

# Open Research Online

---

The Open University's repository of research publications and other research outputs

## End-to-End Ensemble Modelling for Water Resources Planning Under Uncertainty

### Thesis

#### How to cite:

Counsell, Christian John Adam (2018). End-to-End Ensemble Modelling for Water Resources Planning Under Uncertainty. PhD thesis The Open University.

For guidance on citations see [FAQs](#).

© 2017 The Author



<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Version: Version of Record

Link(s) to article on publisher's website:

<http://dx.doi.org/doi:10.21954/ou.ro.0000dcf7>

---

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

---

[oro.open.ac.uk](http://oro.open.ac.uk)

Open University: Earth and Environmental Sciences  
Affiliated Research Centre: HR Wallingford Limited

# End-to-end ensemble modelling for water resources planning under uncertainty

A thesis submitted for the degree of Doctor of Philosophy

Christian John Adam Counsell  
BEng Civil Engineering  
MSc Irrigation Engineering  
June 2017



# Abstract

A hydrological model ensemble, considering two model structures (CatchMOD and PDM), parameter uncertainty and two contrasting methods for estimating potential evapotranspiration (PET), is developed to investigate the relative significance of different sources of uncertainty for water resources planning in the Thames basin. This model ensemble is driven by an ensemble of UKCP09 probabilistic and Future Flows climate change projections, for the 2030s, 2050s and 2080s, to quantify the projected impacts on a range of metrics of relevance to water resources planners using a water resources system model of London.

These sources of supply-side uncertainty are shown to be significant, with the uncertainty associated with the climate change scenarios the largest but hydrological modelling uncertainty, and the method used to estimate PET also shown to be considerable. In terms of overall impacts, the central estimates for the 2030s, 2050s and 2080s are reductions in available resource of around 7%, 11% and 14% respectively. These impacts are shown to equate to economic costs of the order of £360m, £610m and £735m respectively to mitigate such reductions in supply.

The range of uncertainty within each time-horizon is large, greater than the differences between the time-horizons, presenting a significant challenge in deciding the level and timing of investments to mitigate emerging risks. As an example, impacts considered reasonably likely by the 2080s (e.g. a central estimate of 14% impact on deployable output using both PET methods) may be as likely by the 2030s (e. g. using only the modified Penman-Monteith PET method). The estimates of future supply reliability are contrasted with demand forecasts and whilst the pressure associated with the latter is shown to be greater, both are significant and subject to large degrees of uncertainty.

This thesis also highlights the need for detailed examination of hydrological model structures to provide evidence as to their strengths and weaknesses in their representation of key processes, particularly during droughts. The limitations of the climate change products currently used in the industry, particularly with regards to droughts and estimating changes in PET, are also explored.

Significant ongoing research is developing decision-making approaches to support the planning of robust and resilient systems under an uncertain future. This thesis demonstrates that alongside this development, more research is needed to understand, identify and quantify the significant sources of uncertainty that need to be considered as part of the decision-making process.





## Acknowledgements

Firstly, I would like to thank Dr Steven Wade, Dr Richard Whitehouse and HR Wallingford for providing the opportunity to undertake this PhD. I would also like to thank Dr Paul Samuels for providing valuable feedback during the writing-up of this thesis, his positive support throughout and the emphasis he places on academic research at HR Wallingford.

I would like to express my huge gratitude to Professor Jim Freer through whom I have developed significant respect and appreciation of the research process and who has helped me unpick and rebuild my understanding of hydrological modelling and uncertainty. Professor Jim Freer's feedback throughout the research and writing up of the thesis has been greatly appreciated and without which none of what follows would have been possible.

This thesis has also benefitted from the upgrade viva with Professor Hannah Cloke and Professor Neil McIntyre who constructively challenged both what I had done and proposed to deliver. I would also like to acknowledge my examiners for their contributions and the enjoyable and interesting discussions during the viva.

Finally, I must make a huge thank you to my family for providing invaluable support, understanding and sometimes welcome distractions, even as life decided to through some significant challenges in our direction.



# Table of contents

<b>Acknowledgements .....</b>	<b>v</b>
<b>Table of contents .....</b>	<b>vii</b>
<b>List of figures .....</b>	<b>ix</b>
<b>List of tables .....</b>	<b>xviii</b>
<b>Glossary .....</b>	<b>xxi</b>
<b>List of abbreviations .....</b>	<b>xxiii</b>
<b>Chapter 1 Introduction .....</b>	<b>1</b>
1.1.1 The water resources planning process in England and Wales .....	1
1.1.2 Climate change .....	3
1.1.3 Water resources planning under uncertainty .....	4
1.2 <i>Research questions: End-to-end ensemble modelling for water resources planning under uncertainty</i> .....	4
1.3 <i>Structure of this thesis</i> .....	5
<b>Chapter 2 Literature review .....</b>	<b>9</b>
2.1 <i>Hydrological modelling</i> .....	9
2.1.1 Typology of hydrological models .....	9
2.1.2 The components of commonly used conceptual hydrological models .....	11
2.1.3 Hydrological uncertainty and model parameterisation .....	12
2.1.4 The influence of alternative methods for estimating potential evapotranspiration (PET) on model prediction uncertainties .....	17
2.2 <i>Climate change projections for water resources planning in the UK</i> .....	18
2.2.1 Global Climate Models (GCMs) / General Circulation Models .....	19
2.2.2 History of climate change projections in the UK .....	26
2.3 <i>Assessing the hydrological impacts of climate change for water resources planning</i> .....	30
2.3.1 Water Resources industry focussed studies assessing the impacts of climate change .....	35
2.3.2 Research into the climate change impacts for water resources planning .....	37
2.3.3 Chapter summary .....	51
<b>Chapter 3 Methodology .....</b>	<b>55</b>
3.1 <i>Introduction</i> .....	55

3.2	<i>Description of the Thames catchment</i> .....	55
3.3	<i>Development of a hydrological modelling framework for water resources planning</i> .....	56
3.3.1	Choice of hydrological model structures.....	57
3.3.2	The Thames Catchment Model (TCM) .....	59
3.3.3	The Probability Distributed Model (PDM) .....	60
3.4	<i>Data used to drive the hydrological models</i> .....	61
3.4.1	Historical river flow data .....	61
3.4.2	Historical climate data.....	64
3.5	<i>Chapter summary</i> .....	74
<b>Chapter 4 A hydrological model ensemble for water resources planning under uncertainty .....</b>		<b>77</b>
4.1	<i>Introduction</i> .....	77
4.2	<i>Objectives for the research presented in this chapter</i> .....	77
4.3	<i>Development of a Multi-Site Model Ensemble (MSME) for water resources planning</i> .....	78
4.3.1	Uncertainty evaluation.....	78
4.3.2	Incorporating the use of two contrasting methods for estimating PET .....	87
4.3.3	Exploring a gridded soil moisture accounting scheme .....	88
4.3.4	Constructing the Multi-Site Model Ensemble (MSME).....	89
4.4	<i>Assessing the predictive skill of the MSME for water resources planning</i> .....	93
4.4.1	River Thames at Kingston (NRFA catchment 39001) .....	94
4.4.2	Reproduction of flow metrics at other locations in the Thames basin .....	107
4.5	<i>Chapter summary</i> .....	115
<b>Chapter 5 Impacts of climate change on river flows for water resources planning .....</b>		<b>119</b>
5.1	<i>Introduction</i> .....	119
5.2	<i>Objectives</i> .....	119
5.3	<i>Climate change projections for water resource planners</i> .....	120
5.3.1	The ‘probabilistic’ UKCP09 climate change projections.....	120
5.3.2	The transient Future Flows climate projections.....	129
5.3.3	Comparison of changes in precipitation and PET from the Future Flows and UKCP09 projections .....	133
5.4	<i>Hydrological modelling using the UKCP09 and Future Flows climate projections</i> .....	135
5.4.1	Comparison between estimated river flow metrics using the Future Flows projections and the observed climatological record over the 1961 to 2008 period.....	135
5.4.2	Assessing the impacts of climate change on river flows for water resources planning.....	140
5.5	<i>Chapter summary</i> .....	154

<b>Chapter 6</b>	<b>Future reliability of London's public water supplies.....</b>	<b>159</b>
6.1	<i>Introduction.....</i>	159
6.2	<i>Objectives.....</i>	159
6.3	<i>London's water resources system.....</i>	160
6.3.1	London's water resource system .....	160
6.3.2	Water resources and drought planning in London.....	161
6.3.3	Assessing the supply-demand balance.....	162
6.3.4	Modelling of London's water resource system .....	163
6.4	<i>Baseline modelling of the London water resource zone .....</i>	172
6.4.1	Performance of Kestrel – WRM compared to the WARMS system .....	172
6.4.2	Baseline water resources modelling using the MSME.....	174
6.4.3	Hydrological uncertainty related to flows on the River Lee .....	179
6.5	<i>Impacts of climate change on the water resources of London .....</i>	181
6.5.1	Projected impacts of climate change on Levels of Service.....	182
6.5.2	Projected impacts of climate change on Deployable Output .....	189
6.5.3	Uncertainty associated with the projections of the demand for water .....	198
6.6	<i>Chapter summary.....</i>	201
<b>Chapter 7</b>	<b>Conclusions, limitations and recommendations for further research.....</b>	<b>205</b>
7.1	<i>Introduction.....</i>	205
7.2	<i>Summary of research conclusions .....</i>	206
7.3	<i>Limitations of research .....</i>	214
7.4	<i>Suggestions for future research.....</i>	219
7.5	<i>Concluding remarks .....</i>	221
<b>Appendix A</b>	<b>– Inventory of selected conceptual hydrological models.....</b>	<b>249</b>
<b>Appendix B</b>	<b>– Mean monthly changes in PET for the 2030s and 2050s.....</b>	<b>255</b>

## List of figures

<b>Figure 2-1:</b>	<b>Timeline of the publication of climate change scenarios for the UK.....</b>	<b>27</b>
<b>Figure 2-2:</b>	<b>Overview of UKCP09 climate change projections. ....</b>	<b>28</b>
<b>Figure 3-1:</b>	<b>Thames catchment showing the locations considered in this thesis. ....</b>	<b>56</b>

<b>Figure 3-2: The TCM conceptual model structure (semi-distributed form shown) with original CatchMod formulation (reproduced from Cloke et al., 2010) shown in top-left box for reference.....</b>	<b>60</b>
<b>Figure 3-3: The PDM-AQ conceptual model structure (semi-distributed form shown) with original PDM formulation (adapted from Moore, 2007) shown in top-left box for reference. 62</b>	<b>62</b>
<b>Figure 3-4a and b: Comparison of estimated monthly total PET using the Oudin equation (red) and the modified form of the Penman-Monteith equation (black) from 1961 to 1972 (Figure 3-4a, upper plot) and 1973 to 1985 (Figure 3-4b, lower plot) across the Thames basin to Kingston. ....</b>	<b>66</b>
<b>Figure 3-5: Boxplot showing the distribution of winter (October to April), summer (April to September) and annual (January to December) rates of PET (to Kingston) calculated using the Oudin (red) and Penman-Monteith methods (black) between 1961 and 2008. ....</b>	<b>68</b>
<b>Figure 3-6: Rolling 365-day cumulative deficit/surplus PET using both the Oudin (red) and modified Penman-Monteith (black) methods relative to their respective historical long-term average (rolling 365-day catchment average over the Thames catchment to Kingston). ....</b>	<b>69</b>
<b>Figure 3-7: Rolling 365-day cumulative deficit/surplus precipitation relative to the historical long-term average (rolling 365-day catchment average precipitation over the Thames catchment to Kingston). ....</b>	<b>70</b>
<b>Figure 3-8a to c: SPEI (black for modified Penman-Monteith PET method and red for Oudin PET method) and SPI (blue) indices for 12-month (a, top plot), 18-month (b, middle plot) and 24-month (c, lower plot) time windows across the Thames catchment to Kingston. ....</b>	<b>73</b>
<b>Figure 4-1: Distribution of PDM-AQ model performance (top 5% of sampled model parameter sets), as measured by the six water resources relevant performance measures (see Chapter 3), for the Thames catchment to Kingston. Performance scores (not rescaled) calculated for the baseline period 1961 – 2008 using the Oudin PET method. ....</b>	<b>83</b>
<b>Figure 4-2: Distribution of TCM model performance (top 5% of sampled model parameter sets), as measured by the six water resources relevant performance measures (see Chapter 3), for the Thames catchment to Kingston. Performance scores (not rescaled) calculated for the baseline period 1961 – 2008 using the Oudin PET method. ....</b>	<b>84</b>
<b>Figure 4-3: Comparison of model performance, as measured by the six water resources relevant performance measures, for the PDM-AQ HRU catchment ensembles for each study site over all evaluation periods. Models driven using the Oudin (red) and modified Penman-Monteith (black) methods for estimating PET shown. ....</b>	<b>89</b>
<b>Figure 4-4: Comparison of model performance, as measured by the six water resources relevant performance measures, for the PDM-AQ HRU catchment ensembles for each study site over</b>	

all evaluation periods. Only models driven using the Oudin method for estimating PET are shown.....	90
Figure 4-5: Comparison of model performance, as measured by the six water resources relevant performance measures, for the 39001 catchment ensemble and the MSME-OU (models driven using the Oudin method for estimating PET) over all evaluation periods. ....	92
Figure 4-6: Comparison of model performance, as measured by the six performance measures, for the original 39001 catchment ensemble and the MSME-PE (using models driven using the modified Penman-Monteith method for estimating PET) over all evaluation periods. ....	93
Figure 4-7: Simulated FDC (1961 to 2008) at Kingston shown as grey shading (5 <sup>th</sup> to 95 <sup>th</sup> percentile estimates) and black dashed-line (50 <sup>th</sup> percentile estimate). Historical FDC (1961 to 2008) is shown in red. Note that the y-axis is the natural logarithm of flows to emphasise lower flows. ....	94
Figure 4-8: FDC (1971 to 1980) at Kingston. Simulated FDC shown as grey shading (5 <sup>th</sup> to 95 <sup>th</sup> percentile estimates) and black dashed-line (50 <sup>th</sup> percentile estimate). Historical FDC (1971 to 1980) is shown in red. ....	95
Figure 4-9: Simulated FDC (2001 to 2008) at Kingston shown as grey shading (5 <sup>th</sup> to 95 <sup>th</sup> percentile estimates) and black dashed-line (50 <sup>th</sup> percentile estimate). Historical FDC (2001 to 2008) is shown in red. ....	95
Figure 4-10: Simulated MSF from each evaluation period at Kingston. Grey shading presents the ‘unconditioned (i.e. equally weighted)’ distribution and the black box-plots the weighted distribution with ensemble predictions conditioned by all evaluation periods (1961 to 2008 – see Section 3.4). Values from the historical record are shown as red triangles. The years shown on y-axis are the five driest summers in the historical record, ranked in terms of MSF.....	97
Figure 4-11: Simulated hydrograph during 1975 and 1976 at Kingston shown in grey shading (5 <sup>th</sup> to 95 <sup>th</sup> percentile estimates) and black dashed line (50 <sup>th</sup> percentile estimate). Values from the historical record are shown in red. ....	98
Figure 4-12: Simulated hydrograph from 1990 to 1992 at Kingston shown in grey shading and black dashed line. Equivalent values from the historical record shown in red. ....	99
Figure 4-13: Simulated hydrograph from 2004 to 2006 at Kingston shown in grey shading and black dashed line. Equivalent values from the historical record shown in red. ....	101
Figure 4-14: Reproduction of trigger threshold activations at Kingston during Jan 1971 to Dec 1980 shown in red in top plot. Grey shading indicates periods in historical record where trigger threshold is crossed, with the historical flow record itself shown in black on bottom plot. ....	102



<b>Figure 4-15: Reproduction of trigger threshold activations at Kingston during Jan 1981 to Dec 1992 shown in red in top plot. Grey shading indicates periods in historical record where trigger threshold is crossed, with the historical flow record itself shown in black on bottom plot. ....</b>	<b>102</b>
<b>Figure 4-16: Reproduction of trigger threshold activations at Kingston during Jan 1981 to Dec 1992 shown in red in top plot. Grey shading indicates periods in historical record where trigger threshold is crossed, with the historical flow record itself shown in black on bottom plot. ....</b>	<b>103</b>
<b>Figure 4-17: Reproduction of lowest number of days in winter above Q50 from each evaluation period at Kingston. Note that black circles are the full extent of the modelled values. Equivalent values from the historical record shown as red triangles. ....</b>	<b>106</b>
<b>Figure 4-18: Simulated hydrograph for the River Thames at Kingston during 1996 and 1997 shown in grey shading and black dashed line. Equivalent values from the historical record shown in red. ....</b>	<b>107</b>
<b>Figure 4-19: Historical (red) and simulated (black/grey) FDCs for the period 1961 to 2008 for: top plot) River Windrush at Newbridge (NRFA catchment: 39006) and bottom plot) River Wey at Tilford (NRFA catchment: 39011). ....</b>	<b>108</b>
<b>Figure 4-20: Historical (red) and simulated (black/grey) FDCs for: top plot) River Kennet at Theale - from 1963 to 2008 (NRFA catchment 39016) and bottom plot) River Cherwell at Enslow Mill - from 1967 to 2008 (NRFA catchment 39021). ....</b>	<b>109</b>
<b>Figure 4-21: Historical (red) and simulated (black/grey) FDCs for the 1971 to 1980 period for: top plot) River Windrush at Newbridge and bottom plot) River Wey at Tilford.....</b>	<b>110</b>
<b>Figure 4-22: Historical (red) and simulated (black/grey) FDCs for the 1971 to 1980 period for: top plot) River Kennet at Theale and bottom plot) River Cherwell at Enslow Mill. ....</b>	<b>111</b>
<b>Figure 4-23: Reproduction of trigger threshold activations (195 Ml/d, ~2.3 m<sup>3</sup>/s, for the River Kennet at Theale) during Jan 1971 to Dec 1980 shown as a red line in top plot. Grey shading indicates periods in historical record where trigger threshold is crossed, with historical flow record itself shown in black on bottom plot.....</b>	<b>112</b>
<b>Figure 4-24: Reproduction of the flow hydrograph for the River Kennet at Theale during 1975 and 1976. Simulated flows shown as grey shading (5<sup>th</sup> to 95<sup>th</sup> percentile estimate) and black dashed line (50<sup>th</sup> percentile estimate). Equivalent values from the historical record shown in red. ....</b>	<b>113</b>
<b>Figure 4-25: Reproduction of trigger threshold activations (90 Ml/d, ~1.0 m<sup>3</sup>/s) for the River Wey at Tilford) during Jan 1971 to Dec 1980 period shown as a red line in top plot. Grey shading indicates periods in historical record where trigger threshold is crossed, with historical flow record itself shown in black on bottom plot.....</b>	<b>114</b>

<b>Figure 4-26: Reproduction of the flow hydrograph for the River Wey at Tilford during 1975 and 1976. Simulated flows shown as grey shading (5<sup>th</sup> to 95<sup>th</sup> percentile estimates) and black dashed line (50<sup>th</sup> percentile estimate). Values from the historical record are shown in red.</b>	<b>114</b>
<b>Figure 5-1: Seasonal change (percentage) in mean daily precipitation and temperature (in degrees Celsius) from the UKCP09 2030 Medium Emission scenario for the full 10,000 member ensemble and the 20-member sub-sample. Changes are relative to a 1961 to 1990 baseline period.....</b>	<b>122</b>
<b>Figure 5-2: Seasonal change (percentage) in mean daily precipitation and temperature (in degrees Celsius) from the UKCP09 2050 Medium Emission scenario for the full 10,000 member ensemble and the 20-member sub-sample. Changes are relative to a 1961 to 1990 baseline period.....</b>	<b>123</b>
<b>Figure 5-3: Seasonal change (percentage) in mean daily precipitation and temperature (in degrees Celsius) from the UKCP09 2080 Medium Emission scenario for the full 10,000 member ensemble and the 20-member sub-sample. Changes are relative to a 1961 to 1990 baseline period.....</b>	<b>124</b>
<b>Figure 5-4: Mean monthly change (as a percentage) in precipitation (top plot) and PET (bottom plot, using the Oudin method) from the UKCP09 2030, 2050 and 2080 Medium Emission scenarios. Values shown are for the 10<sup>th</sup>, 25<sup>th</sup>, 50<sup>th</sup>, 75<sup>th</sup> and 90<sup>th</sup> percentiles for the full 10,000 member ensembles and each selected scenario from the 20-member sub-samples. Changes are relative to a 1961 to 1990 baseline. Each y-axis provides the UKCP09 scenario ID for each sampled scenario for each future time-horizon.....</b>	<b>126</b>
<b>Figure 5-5: Comparison of mean monthly changes in PET under the Oudin (red dashed lines) and modified Penman-Monteith (black dashed lines) methods and the mean monthly change in mean daily temperature (green dashed lines) for 20 UKCP09 sampled projections for the 2080s under a Medium Emission scenario. Changes are relative to a 1961 to 1990 baseline. ....</b>	<b>127</b>
<b>Figure 5-6: Comparison of the distribution of mean monthly changes in PET under the Oudin (red) and modified Penman-Monteith (black) methods for 20 UKCP09 sampled projections for the 2080s under a Medium Emission scenario. Boxplot whiskers extend to the 5<sup>th</sup> and 95<sup>th</sup> percentiles of the 20 sampled scenarios. Changes are relative to a 1961 to 1990 baseline. ....</b>	<b>128</b>
<b>Figure 5-7: Box plots of monthly precipitation totals across the Thames basin from the historical record (E-Obs v 3.0 – see 3.4.2) as the wider black box-plots and the eleven-member Future Flows (FF) ensemble over the 1961 to 2008 time period (with each individual scenario shown as a light grey box plot). ....</b>	<b>129</b>
<b>Figure 5-8: Box plots of monthly PET totals across the Thames basin (from the historical record (based on the modified Penman-Monteith method and Met Office gridded historical</b>	

weather data – see Section 3.4.2) as the wider black box-plots and the eleven-member FF ensemble over the 1961 to 2008 time period (with each individual scenario shown as a light grey box plot).....	130
<b>Figure 5-9: Rolling 365-day deficit/surplus precipitation across the Thames basin from the eleven FF ensemble members relative to the historical LTA 365-day precipitation (E-Obs v 3.0). .....</b>	<b>131</b>
<b>Figure 5-10: Monthly change (as a percentage) in mean monthly precipitation (top plot) and PET (bottom plot, using the modified Penman-Monteith method) from the UKCP09 2030, 2050 and 2080 Medium Emission scenarios - 10<sup>th</sup>, 25<sup>th</sup>, 50<sup>th</sup>, 75<sup>th</sup> and 90<sup>th</sup> percentiles and each scenario from the eleven member FF ensemble. Changes are relative to a 1961 to 1990 baseline. ....</b>	<b>134</b>
<b>Figure 5-11: MSME-PM simulated FDC for Kingston for the 1961 to 2008 climatology from each FF ensemble member (50<sup>th</sup> percentile for each FF scenario shown as a distinct colour) and the historical 1961 to 2008 climatology (shown as grey shared area [5<sup>th</sup> to 95<sup>th</sup> percentile] and black triangles [50<sup>th</sup> percentile]). Values from the historical flow record (1961 to 2008) shown as black crosses.....</b>	<b>139</b>
<b>Figure 5-12: Lowest MSF at Kingston for the lowest years in the 1961 to 2008 historical record (black triangles). Distribution of simulated values (black box-plots) and the 50<sup>th</sup> percentile MSME-PM simulated value for the five lowest MSF values from each FF ensemble member (each FF scenario shown as a distinct colour) also shown.....</b>	<b>141</b>
<b>Figure 5-13: The range of predicted changes at top) Q95, middle) Q70 and bottom) Q50 at under each FF scenario (MSME-PM simulations) and the UKCP09 20-member sub-samples (MSME simulations) for the 2030s (red), 2050s (green) and 2080s (blue) relative to a 1961-1990 baseline. ....</b>	<b>144</b>
<b>Figure 5-14: Change (as a percentage) in top) Q95 and bottom) Q50 at Kingston for individual UKCP09 scenario members for the 2080s (relative to a 1961-1990 baseline) using the MSME-OU (red) and MSME-PM (green). ....</b>	<b>145</b>
<b>Figure 5-15: Historical lowest MSF at Kingston for the four lowest years in the 1961 to 1990 historical record (black triangles) along with the distribution of simulated values driven by the historical climatology (black box-plots). The equivalent sequences from the UKCP09 sub-sample (MSME simulations, shown as black box-plots) and FF (MSME-PM simulations, 50<sup>th</sup> percentile only shown for each FF scenario as solid circle) for the 2030s (blue), 2050s (red) and 2080s (green). ....</b>	<b>147</b>
<b>Figure 5-16: Comparison between MSME-OU (solid shading) and MSME-PM (no shading) simulations for the UKCP09 20-member sub-samples in terms of the impact on the lowest MSF event (the perturbed summer of 1976) for the 2030s (blue), 2050s (red) and 2080s (green). ....</b>	<b>148</b>

Figure 5-17: MSME-PM simulated hydrograph at Kingston for 2080 to 2081 for the ‘ <i>afixi</i> ’ FF scenario. The blue and red shaded areas (and dotted red line when covered by the modelled daily flow percentiles) show the extent of the historical envelope of daily maximum and minimum flows during the historical flow record 1961 to 1990. ....	150
Figure 5-18: MSME-PM simulated hydrograph at Kingston for 2067 to 2068 for the ‘ <i>afixl</i> ’ FF scenario. The blue and red shaded areas (and dotted red line when covered by the modelled daily flow percentiles) show the extent of the historical envelope of daily maximum and minimum flows during the historical flow record 1961 to 1990. ....	151
Figure 5-19: Comparison of the MSME simulated impacts on Q95 (top), Q70 (middle) and Q50 (bottom) for different sub-catchments using the UKCP09 climate projections for 2030s (black box-plot), 2050s (red box-plot) and 2080s (green box-plot), all relative to a 1961 to 1990 baseline under a Medium Emission scenario. Catchment BFI (Gustard et al., 1992) also presented. ....	152
Figure 5-20: Comparison of the simulated impacts on Q95 (top), Q70 (middle) and Q50 (bottom) for different sub-catchments using the UKCP09 climate projections for the 2080s (relative to a 1961 to 1990 baseline) under a Medium Emission scenario, with results presented for MSME-OU (black box-plot), MSME-PM (red box-plot) and MSME (green box-plot) simulations respectively. Catchment BFI (Gustard et al., 1992) also presented. ....	153
Figure 6-1: Overview of the Lower Thames system including locations of storage and surface water intakes (Thames Water, 2014). ....	161
Figure 6-2: Drought severity as a function of naturalised flow at Teddington (April to September) between 1900 and 2006 (taken from Thames Water’s Final Drought Plan, 2013). ....	162
Figure 6-3: Conceptual representation of the Kestrel – WRM model of the London water resources system. ....	165
Figure 6-4: Lower Thames Operating Agreement Control Diagram (LTOACD) showing demand reduction thresholds (Levels 1 to 4 as dashed lines) and minimum environmental flow rates as a function of London’s aggregate reservoir storage (LAS). ....	167
Figure 6-5: Monthly demand factors for supply-demand modelling (Thames Water, 2014). ....	171
Figure 6-6: Profile of the LAS during selected periods in the historical climate record using the WRMP demand profile and existing water resources system modelled using the WARMS model (Thames Water, 2014). The black dashed lines and coloured shading reflect the LTOACD (see Figure 6-4). ....	173
Figure 6-7: Profile of the LAS during the 1975 to 1976 period using the Kestrel - WRM model (solid cyan line). The black dashed lines and red/yellow shading reflect the LTOACD (see	

Figure 6-4). The yellow shaded areas are Levels 1 and 2, with the red areas Levels 3 and 4 – as used in Figure 6-6.....	174
Figure 6-8: Profile of LAS during the 1975 to 1976 period comparing the MSME prediction (black line and grey shading) and the profile generated using the historical flow record (red line). The coloured areas reflect the LTOACD (Levels 1 to 4, green to red shading – see Figure 6-4).....	175
Figure 6-9: Forecast LTOA levels during the 1975 to 1976 period comparing the MSME prediction (bottom plot) and the profile generated using the historical flow record (top plot). The coloured areas reflect the LTOACD (Levels 1 to 4, green to red shading – see Figure 6-4).....	175
Figure 6-10: Profile of the LAS during the 2001 to 2008 period comparing the MSME predictions (black line and grey shading) and the profile generated using the historical flow record (red line). The coloured areas reflect the LTOACD (Levels 1 to 4, green to red shading – see Figure 6-4).....	176
Figure 6-11: Profile of the LAS during the 1975 to 1976 period comparing the predictions using the MSME –OU ensemble (black line and grey shading) against the profile generated using the historical flow record (red line). The coloured areas reflect the LTOACD (Levels 1 to 4, green to red shading – see Figure 6-4). ....	177
Figure 6-12: Profile of the LAS during the 1975 to 1976 period comparing the predictions using the MSME-PM ensemble (black line and grey shading) against the profile generated using the historical flow record (red line). The coloured areas reflect the LTOACD (Levels 1 to 4, green to red shading – see Figure 6-4). ....	178
Figure 6-13: Profile of the LAS during the 1975 to 1976 period from the historical climate record. Range of values for reservoir drawdown using the MSME-200 is shown as a black line and grey shading. The storage profile using the historical flow record is shown as a red dashed line for reference. The coloured areas reflect the LTOACD (Levels 1 to 4, green to red shading – see Figure 6-4). ....	179
Figure 6-14: Profile of the LAS during the 1975 to 1976 period from the historical climate record. The range of values for reservoir drawdown using the MSME-200-LEE is shown as a black line and grey shading. The storage profile using the historical flow record is shown as a red dashed line for reference. The coloured areas reflect the LTOACD (Levels 1 to 4, green to red shading – see Figure 6-4). ....	181
Figure 6-15: Histogram of the frequency of Level 4 LTOACD transgressions for each UKCP09 climate change scenario sub-sample members for the 2030s (top), 2050s (middle) and 2080s (bottom). The transgressions for the baseline climatology and combined 20-member ensemble are represented by the small circles and small squares above the histograms respectively.....	184

<b>Figure 6-16: Forecast frequency of years during which a LTOACD Level 4 breach is realised for different time-horizons and distinguishing between the different approaches to estimating PET. ....</b>	<b>185</b>
<b>Figure 6-17: The distribution of the number of LTOACD Level 4 days (breaches) for the for each UKCP09 climate change scenario sub-sample member for the 2030s. The distribution for the baseline climatology and combined 20-member ensemble are presented at the far left and far right of each plot respectively. ....</b>	<b>186</b>
<b>Figure 6-18: The distribution of the number of LTOACD Level 4 days (breaches) for the for each UKCP09 climate change scenario sub-sample member for the 2050s. The distribution for the baseline climatology and combined 20-member ensemble are presented at the far left and far right of each plot respectively. ....</b>	<b>187</b>
<b>Figure 6-19: The distribution of the number of LTOACD Level 4 days (breaches) for the for each UKCP09 climate change scenario sub-sample member for the 2080s. The distribution for the baseline climatology and combined 20-member ensemble are presented at the far left and far right of each plot respectively. ....</b>	<b>188</b>
<b>Figure 6-20: The distribution of the DO for each UKCP09 climate change scenario sub-sample member for the 2030s. Calculated as the DO at which system performance is at least as good as for the baseline condition. The distribution for the combined 20-member ensembles is presented at the far left and far right of each plot respectively. Baseline DO is 2,144 Ml/d (~24.8 m<sup>3</sup>/s). ....</b>	<b>191</b>
<b>Figure 6-21: The distribution of the DO for each UKCP09 climate change scenario sub-sample member for the 2050s. Calculated as the DO at which system performance is at least as good as for the baseline condition. The distribution for the combined 20-member ensembles are presented at the far left and far right of each plot respectively. Baseline DO is 2,144 Ml/d (~24.8 m<sup>3</sup>/s). ....</b>	<b>192</b>
<b>Figure 6-22: The distribution of the DO for each UKCP09 climate change scenario sub-sample member for the 2080s. Calculated as the DO at which system performance is at least as good as for the baseline condition. The distribution for the combined 20-member ensembles is presented at the far left and far right of each plot respectively. Baseline DO is 2,144 Ml/d (~24.8 m<sup>3</sup>/s). ....</b>	<b>193</b>
<b>Figure 6-23: The distribution of the DO for each UKCP09 sub-sample ensemble for the 2030s, 2050s and 2080s. Calculated as the DO at which system performance is at least as good as for the baseline condition. Baseline DO is 2,144 Ml/d (~24.8m<sup>3</sup>/s). ....</b>	<b>194</b>
<b>Figure B. 1: Comparison of mean monthly changes in PET under the Oudin (red dashed lines) and modified Penman-Monteith (black dashed lines) methods and the mean monthly change in mean daily temperature (green dashed lines) for 20 UKCP09 sampled projections for the 2030s under a Medium Emission scenario. Changes are relative to a 1961 to 1990 baseline. ....</b>	<b>257</b>

<b>Figure B. 2: Comparison of mean monthly changes in PET under the Oudin (red dashed lines) and modified Penman-Monteith (black dashed lines) methods and the mean monthly change in mean daily temperature (green dashed lines) for 20 UKCP09 sampled projections for the 2050s under a Medium Emission scenario. Changes are relative to a 1961 to 1990 baseline. ....</b>	<b>258</b>
---	------------

## List of tables

<b>Table 2-1: Selected research studies investigating the hydrological impact of climate change (with an emphasis on UK studies of relevance to the water resources planning industry).....</b>	<b>31</b>
<b>Table 3-1: Hydrogeological classifications (after Bloomfield et al., 2009) for the modelled catchments. ....</b>	<b>59</b>
<b>Table 3-2: Lowest historical mean summer flow (April to September, m<sup>3</sup>/s) for five separate time-periods for five locations within the Thames basin. ....</b>	<b>63</b>
<b>Table 3-3: Selected flow percentiles from the historical flow record at Kingston for five separate time-periods (calendar years).....</b>	<b>64</b>
<b>Table 3-4: Sources of historical climate data used in this thesis. ....</b>	<b>64</b>
<b>Table 3-5: Comparison of mean annual precipitation and the frequency and severity of dry periods in the Thames catchment to Kingston for different periods in the historical record. .</b>	<b>70</b>
<b>Table 4-1: Metrics used to evaluate model performance. ....</b>	<b>80</b>
<b>Table 4-2: Number of models (out of 10,000 LHS parameter sets) that achieved a non-zero score during each of the five separate simulation time-periods. Models driven using the Oudin (Oudin et al., 2005) method for estimating PET only are presented. Catchment locations are shown in Figure 3-1. ....</b>	<b>86</b>
<b>Table 4-3: Remaining number of models (out of 10,000 LHS parameter sets) achieving a non-zero performance score as the cumulative information content from each simulation period is added in chronological order – models driven using the Oudin method for estimating PET only are presented. ....</b>	<b>86</b>
<b>Table 4-4: Comparison of the number of retained models (out of 10,000 LHS parameter sets), considering all evaluation periods using the two different approaches for estimating PET....</b>	<b>88</b>
<b>Table 4-5: Comparison of the number of models retained (out of 10,000 LHS parameter sets) using the MSME with the '39001 catchment ensemble' for the Thames catchment to Kingston. ....</b>	<b>92</b>
<b>Table 4-6: Forecasting skill measures used to assess the performance of the MSME in identifying periods when trigger conditions are satisfied. ....</b>	<b>104</b>

<b>Table 4-7: MSME skill in predicting periods when the trigger condition is satisfied in the historical record. ....</b>	<b>105</b>
<b>Table 5-1: Comparison between the historical precipitation record and each FF ensemble member across for the 1961 to 2008 period for significant precipitation deficit periods. Mean and co-efficient of variance (CoV) across the FF ensemble member values are also shown. Notably dry events are these including a deficit of greater than 100mm from LTA. .</b>	<b>132</b>
<b>Table 5-2: Comparison between mean historical PET and the mean from each FF ensemble member for the 1961 to 2008 period. Mean and co-efficient of variance (CoV) across the FF ensemble member mean values are also shown. ....</b>	<b>132</b>
<b>Table 5-3: Combined Overlap Percentage scores for each FF ensemble member compared to that using the historical climatology. Score based on selected flow percentiles (Q1, Q5, Q10, Q30, Q50, Q70, Q90, Q95 and Q99) over the 1961 to 2008 period. ....</b>	<b>137</b>
<b>Table 5-4: 5<sup>th</sup>, 50<sup>th</sup> and 95<sup>th</sup> predicted change (percentage relative to a 1961 to 1990 baseline) using the MSME-PM (Future Flows) and MSME (UKCP09) in Q95, Q70 and Q50 for the Thames at Kingston under the UKCP09 sub-sample and each FF ensemble member for the 2080s.....</b>	<b>142</b>
<b>Table 5-5: Potential events of interest for water resources planning in London in the FF ensemble.....</b>	<b>149</b>
<b>Table 6-1: Estimates for the impacts of demand reduction for London WRZ, expressed as percentages (Thames Water Drought Plan, 2013). ....</b>	<b>171</b>
<b>Table 6-2: Number of days at each LTOACD Level predicted using the MSME compared to those generated using the historical flow sequence, during 1975 to 1976, at Teddington.....</b>	<b>176</b>
<b>Table 6-3: Number of days at each LTOACD Level predicted using the MSME compared to those generated using the historical flow sequence, during 1975 to 1976, at Teddington.....</b>	<b>177</b>
<b>Table 6-4: Number of days at each LTOACD Level predicted using the MSME and the 200 member MSME-200, during 1975 to 1976, at Teddington.....</b>	<b>179</b>
<b>Table 6-5: Number of days at each LTOACD Level predicted using the MSME-200-LEE during 1975 to 1976. ....</b>	<b>180</b>
<b>Table 6-6: Simulated number of years at LTOACD Levels 3 and 4 for the baseline (1961 to 1990) period.....</b>	<b>182</b>
<b>Table 6-7: Simulated number of days at LTOACD Levels 3 and 4 for the baseline (1961 to 1990) period.....</b>	<b>182</b>
<b>Table 6-8: Summary of influence on Deployable Output impacts by time-horizon of different sources of uncertainty. Baseline DO is 2,144 Ml/d (~24.8 m3/s).....</b>	<b>195</b>



<b>Table 6-9: Percentage increases in population in London for the 2050s and 2080s (HR Wallingford Limited, 2015). .....</b>	<b>199</b>
<b>Table 6-10: Household and non-household consumption and leakage assumptions used to develop projections of demand for the CCRA2 water availability report (HR Wallingford Limited, 2015). .....</b>	<b>199</b>
<b>Table 6-11: Percentage increases in the demand for public water supplies in London for the 2050s and 2080s (HR Wallingford Limited, 2015) based on population forecasts and the assumptions detailed in Table 6-10 with regards to per capita consumption, non-household demand and leakage. ....</b>	<b>200</b>

## Glossary

**Deployable Output (DO)** - is the amount of water that can be pumped from a water company's sources (surface and groundwater), constrained by licence, hydrology or hydrogeological factors and works capacity.

**Distribution Input (DI)** - is the average amount of potable water entering the distribution system within the area of supply. It is the sum of the total demand for water from people (i.e. household and non-household use), plus losses from leakage.

**Flow Duration Curve (FDC)** - A flow duration curve is a graphical representation of a ranking of all the flows in a given period, from the lowest to the highest, where the rank is the percentage of time the flow value is equalled or exceeded.

**Levels of Service (LOS)** - describe the average frequency that a company will apply restrictions on water use to customers.

**Per Capita Consumption (PCC)** - Measure of average water use for each person in an appointed water company's area.

**Reliability** – The degree to which a water resource system satisfies its specified Levels of Service, as set-out in the Water Resources Management Plan

**Supply-Demand Balance (SDB)** - The difference between the water available for supply and the demand for water.

**Target Headroom**- represents the minimum buffer that water companies should plan to maintain between supply and demand for water in order to cater for current and future uncertainties.

**Water Available For Use (WAFU)** - The Deployable Output plus bulk supply imports, less bulk supply exports and less reductions made for outage allowance and operational losses.

**Water Resource Zone (WRZ)** - The largest possible zone in which all water resources, excluding external transfers, can be shared. Hence, it is the zone in which all customers are considered experience the same risk of supply failure from a resource shortfall.



## List of abbreviations

<b><i>AET</i></b>	Actual Evapotranspiration
<b><i>AR4</i></b>	IPCC Fourth Assessment Report
<b><i>ASC</i></b>	Adaptation Sub-Committee
<b><i>BFI</i></b>	Base Flow Index
<b><i>BSS</i></b>	Brier Skill Score
<b><i>CAMS</i></b>	Catchment Abstraction Management Strategy
<b><i>CCC</i></b>	Committee on Climate Change
<b><i>CCRA2</i></b>	The second Climate Change Risk Assessment for the UK
<b><i>CHARS</i></b>	Chingford Artificial Recharge Scheme
<b><i>COP</i></b>	Combined Overlap Percentage
<b><i>CoV</i></b>	Co-efficient of Variance
<b><i>CMIP3</i></b>	Third Coupled Model Inter-comparison Project
<b><i>CMIP5</i></b>	Fifth Coupled Model Inter-comparison Project
<b><i>DI</i></b>	Distribution Input
<b><i>DO</i></b>	Deployable Output
<b><i>DP</i></b>	Drought Plan
<b><i>EU</i></b>	European Union
<b><i>FF</i></b>	Future Flows projections
<b><i>FDC</i></b>	Flow Duration Curve
<b><i>GCM</i></b>	General Circulation Model or Global Climate Model
<b><i>GLUE</i></b>	Generalised Likelihood Uncertainty Estimation
<b><i>HRU</i></b>	Hydrological Response Unit

<b>IPCC</b>	Intergovernmental Panel on Climate Change
<b>IRZ</b>	Integrated Resource Zone
<b>LAS</b>	London Aggregate (Reservoir) Storage
<b>LHS</b>	Latin Hypercube Sampling
<b>LogNSE</b>	Nash Sutcliffe Efficiency (on the natural log of flows)
<b>LOS</b>	Levels Of Service
<b>LTA</b>	Long Term Average
<b>LTOA</b>	Lower Thames Operating Agreement
<b>LTOACD</b>	Lower Thames Operating Agreement Control Diagram
<b>MAF</b>	Mean Annual Flow
<b>ME</b>	Medium Emission Scenario
<b>ML/d</b>	Mega Litres per Day (= metres cubed per second *24*60*60 / 1000)
<b>MSF</b>	Mean Summer Flow
<b>MSME</b>	Multi-Site Model Ensemble
<b>MSME-OU</b>	MSME ensemble members derived using the Oudin method for estimating potential evapotranspiration only
<b>MSME-PM</b>	MSME ensemble members derived using the modified Penman-Monteith method for estimating potential evapotranspiration only
<b>MSME-200</b>	200 member MSME sub-sample
<b>MSME-200-LEE</b>	200 member MSME sub-sample water resource simulations combining models for the River Thames and River Lee
<b>MSME-200-LEE-OU</b>	MSME-200-LEE members derived using the Oudin method for estimating potential evapotranspiration only
<b>MSME-200-LEE-PM</b>	MSME-200-LE members derived using the modified Penman-Monteith method for estimating potential evapotranspiration only

<b><i>NHH</i></b>	Non-Household demand
<b><i>NLARS</i></b>	North London Artificial Recharge Scheme
<b><i>NRFA</i></b>	National River Flow Archive
<b><i>NPV</i></b>	Net Present Value
<b><i>NSE</i></b>	Nash Sutcliffe Efficiency
<b><i>ONS</i></b>	Office of National Statistics
<b><i>PBIAS</i></b>	Percent Bias
<b><i>PCC</i></b>	Daily Per Capita Consumption of water
<b><i>PET</i></b>	Potential Evapotranspiration
<b><i>PDM</i></b>	Probability Distributed Moisture Model
<b><i>PDM-AQ</i></b>	Probability Distributed Moisture Model – modified form for groundwater dominated catchments
<b><i>PWS</i></b>	Public Water Supply
<b><i>RCM</i></b>	Regional Climate Model
<b><i>RCP</i></b>	Representative Concentration Pathways
<b><i>RFAT</i></b>	Relative Forecasting Accuracy of Trigger
<b><i>RFFAT</i></b>	Relative False Forecasting Accuracy of Trigger
<b><i>RQ</i></b>	Research Question
<b><i>SCE</i></b>	Shuffled Complex Evolution
<b><i>SDB</i></b>	Supply-Demand Balance
<b><i>SRES</i></b>	Special Report on Emissions Scenarios
<b><i>SLARS</i></b>	South London Artificial Recharge Scheme
<b><i>SPEI</i></b>	Standardized Precipitation and Evapotranspiration Index

<b><i>SPI</i></b>	Standardized Precipitation Index
<b><i>TCM</i></b>	Thames Catchment model (aka CatchMod)
<b><i>TGWTW</i></b>	Thames Gateway Water Treatment Works (desalination plant)
<b><i>TUB</i></b>	Temporary Use Ban
<b><i>TWUL</i></b>	Thames Water Utilities Limited
<b><i>UKCP09</i></b>	UK Climate Projections 2009
<b><i>UKWIR</i></b>	UK Water Industry Research
<b><i>WAFU</i></b>	Water Available For Use
<b><i>WARMS</i></b>	Water Resources Management System
<b><i>WMO</i></b>	World Meteorological Organization
<b><i>WRMP</i></b>	Water Resources Management Plan
<b><i>WRZ</i></b>	Water Resource Zone
<b><i>WTW</i></b>	Water Treatment Works
<b><i>WBGW</i></b>	West Berkshire Ground Water Scheme

# Chapter 1 Introduction

Approximately 2.5% of the world's water is estimated to be available as freshwater with the majority of this freshwater stored in the form of glaciers, ice caps, and hard to access groundwater. It is estimated that around 1.2% of this freshwater is available to meet human needs (Shiklomanov, 1993; ICA, 2012) and that over 1 billion of the world's population are living in areas exposed to water scarcity (e.g. Hayashi et al. 2010; Gosling & Arnell, 2013).

Water is considered to be one of the main mechanisms for people to experience climate change (Jiménez Cisneros et al., 2014). As a result of climate change and population growth, the number of people estimated to become exposed to water scarcity is projected to increase sharply in the future (e.g. Hayashi et al. 2010; Schewe et al. 2014; Gosling & Arnell, 2016). One study (Schewe et al., 2014) has estimated that a global warming of 2°C (from present) could result in approximately an additional 15% of the global population facing a severe decrease (defined as a reduction by more than 20% and a reduction by more than one standard deviation of the 1980 to 2010 annual discharge) in water resources, and the number of people living under absolute water scarcity (defined as less than 500 m<sup>3</sup> per capita per year) by another 40%, compared with the effect of population growth alone.

In the UK, a recent report highlighted that, over the next 30 years, there is likely to be increasing pressures on water availability from rising population, associated development and the impacts of climate change (HR Wallingford, 2015). Further, environmental pressures affecting how much and when water can be abstracted also need to be considered, with the *EU Water Framework Directive* (Directive: 2000/60/EC) setting a number of objectives including ensuring that all water bodies reached “good” ecological status by 2015, and the *EU Habitats Directive* (Directive: 92/43/EEC) requiring Member States to take measures to maintain or restore natural habitats and wild species at a “favourable” conservation status. These environmental objectives are leading to additional restrictions and constraints being attached to some abstraction licenses which will further increase pressure on maintaining adequate water supplies for societal use in the future.

## 1.1.1 The water resources planning process in England and Wales

Annual average rainfall over England and Wales is 890 mm with approximately half of this estimated to be lost by evaporation leaving an average of 465 mm for runoff to rivers or percolation to groundwater. The network of rivers and aquifers provide the vast majority of the water supplied with water companies abstracting almost half of the total amount taken from non-



tidal waters in England and Wales – with over 70 per cent returned as treated effluent (Environment Agency, 2008).

In England and Wales, for which the Environment Agency is the statutory authority, the operative unit for the management of abstractions and discharges is the Catchment Abstraction Management Strategy (CAMS) area, of which 127 cover England and Wales. CAMS are designed to contribute to the implementation of the Water Framework Directive and set out how the water resources of a catchment will be managed including how much freshwater resource is reliably available, how much water the environment needs and the amount of water already licensed for abstraction (Environment Agency, 2010)

Under the *Water Industry Act 1991* (section 37A-37d), brought in by the *Water Act 2003*, water companies have a statutory duty to “prepare, consult, publish and maintain” a Water Resources Management Plan (WRMP) that considers the availability and demand for water. The *Water Resources Management Plan Regulation 2007* (Environment Agency, 2012) sets out the steps a statutory water undertaker must follow with respect to the publication and consultation of a draft water resources management plan, and the publication of its final plan. The *Water Resources Management Plan Direction 2007* details what information the water resources plans should include and provides statutory timescales for this process.

The Environment Agency, with support from Natural Resources Wales, OFWAT, Water UK and the water industry, has developed guidance (Environment Agency, 2012; 2017) that provides a framework for developing and maintaining water resources management plans. This states that the WRMP should cover at least the statutory minimum period of 25 years (the WRMP planning horizon) and include how a water company plans to contribute to the delivery of Water Framework Directive objectives (Environment Agency, 2017).

The WRMP should demonstrate how an efficient, sustainable secure supply of water will be maintained to customers. As part of this, water companies typically use hydrological, hydrogeological and water resource system models, driven by projections of climate change, to quantify resource availability for each year in the planning horizon. These estimates are then assessed against future forecasts of the demand for water. Where deficits are identified, water companies develop a portfolio of secure and sustainable options designed to cost-effectively maintain supplies into the future whilst fulfilling their Water Framework Directive obligations.

### 1.1.2 *Climate change*

Warming of the climate system is unequivocal (e.g. Solomon et al., 2007; IPCC, 2013; Kendon et al., 2016) and the impacts are already being observed in the UK, with the Central England Temperature records indicating a rise of approximately one degree Celsius since the 1970s, with human activity cited as a significant influence (Jenkins et al., 2009). Also, whilst annual mean precipitation over England and Wales has not changed significantly since records began in 1766, rainfall amounts appears to have decreased in summer and increased in winter, although with little change in the latter over the last 50 years (Jenkins et al., 2009). However, recent studies (e.g. Watts et al., 2015; Hannaford, 2015) have concluded that there is currently insufficient evidence across the UK to clearly and consistently demonstrate a link between anthropogenic climate change and changes in precipitation, river flows and groundwater levels.

In the UK, it is anticipated that climate change is very likely to have a significant impact on water resources across the United Kingdom into the future (e.g. Wilby, 2005; Whitehead et al., 2009; Cloke et al., 2010; Christerson et al., 2012; Charlton & Arnell, 2014; HR Wallingford, 2015; Watts et al., 2015). Any significant reduction in water availability, particularly in those areas of the country considered to be already “water stressed” (Environment Agency, 2008; 2013), including London which is considered to be “seriously water stressed”, could have major implications for quality of life. Furthermore, as well as directly impacting the availability of water to meet demand, lower minimum flows imply less volume for dilution and, hence, higher pollution concentrations downstream of point discharges, such as wastewater treatment works (Whitehead et al., 2009). Adaptation to such impacts in the context of water resources may require complex and potentially controversial measures to be implemented at an appropriate time.

The potential impacts of climate change have been considered explicitly in Water Resources Management Plans since 1999. For the next set of WRMPs (2019), two climate change products recommended for use (Environment Agency, 2012; 2017) are the UKCP09 ‘probabilistic’ projections (Murphy et al., 2009b) and the eleven-member Future Flows transient climate projections (Prudhomme et al., 2012). These climate change projections are typically applied to hydrological models to estimate the impacts on river flows which are then input into water resource system models to quantify the impacts on supply system performance.

It is recognised that using hydrological models driven by climate scenarios to project future water availability is subject to a range of uncertainties. These uncertainties include those associated with the hydrological models, their input data and calibration (e.g. Beven et al., 2008; Velázquez et al., 2013; Charlton & Arnell, 2014; Borgomeo et al., 2014; Westerberg et al., 2016) and that

related to the provision and downscaling of climate change projections (e.g. Stainforth et al. 2005; Murphy et al., 2009b; Turner et al., 2014; Watts et al., 2015; Walsh et al., 2015).

### *1.1.3 Water resources planning under uncertainty*

Uncertainty in future supply-side and demand-side conditions along with changing stakeholder priorities are encouraging the development of new methods for water resources planning (Matrosov et al., 2013; Korteling et al., 2013; Huskova et al., 2016; UKWIR, 2016b; UKWIR 2016c). Such decision-making approaches require defensible estimates of the uncertainty associated with all aspects of the decision-making process including predicting future resource availability using hydrological models and climate projections.

Further, the UK industry is moving towards the use of synthetic and stochastic weather sequences for better understanding water resource system resilience (Southern Water, 2014; UKWIR, 2016c) rather than simply relying upon the historical record. This places even greater reliance on the 'skill' of hydrological models, particularly their adequacy to model conditions beyond that to which they have been evaluated. Whatever decision-making approaches are adopted in the industry the treatment of uncertainty will not be completely exhaustive. Faced with current computational constraints evidence is required to understand what are the dominant sources of uncertainty and their potential significance on the performance of water resource systems (i.e. the ability to meet specified levels of service to customers). This evidence needs to be supported by the development of pragmatic approaches to considering and quantifying uncertainty that can be readily applied by the UK water resources industry.

## **1.2 Research questions: End-to-end ensemble modelling for water resources planning under uncertainty**

This thesis contributes to the evidence base for the design and implementation of ensemble modelling studies used as part of the water resources planning process and, in particular, explores the following research questions (RQs):

- RQ1.** How important are the uncertainties associated with hydrological model structures, model parameterisation and the method used to estimate PET and what are their influences on the flow characteristics of relevance to the water resources planning process (Chapter 4)?
- RQ2.** How can the climate projections currently recommended for use by water resource planners in the UK be used as part of an ensemble modelling study and how appropriate

are these projections for water resources planning? How significant is the method used to estimate PET as part of a climate change impact assessment (Chapter 5)?

- RQ3.** How do the uncertainties associated with hydrological modelling, estimating PET and climate change all compare in terms of their potential influence on water resources planning, and how do they compare to demand-side uncertainties? How can the relative significance of different uncertainty sources be quantified in the language of the water resources planner (Chapter 6)?

The Thames catchment and London's water supply system are used as case studies throughout this thesis and further case-study specific research questions (RQs) are posed by this research. These are:

- RQ4.** How do different hydrological models differ in their reproduction of observed river flows and the specific metrics of interest to water resources planners across the Thames catchment and how does this reproduction vary over different time-periods (Chapter 4)?
- RQ5.** What is the projected magnitude of the impacts of climate change on river flows across the Thames catchment and what is the trajectory of these changes beyond the industry standard 25 year planning horizon (Chapter 5)?
- RQ6.** How significant are the different sources of hydrological modelling uncertainty on the performance of London's water resources system? How do the projected impacts of climate change on London's water supplies compare to those anticipated to be caused by population growth and the resulting changes in the demand for water? How do these vary into the future, considering three future time-periods, the 2030s, 2050s and 2080s (Chapter 6)?

### 1.3 Structure of this thesis

Chapter 2 provides a review of relevant academic research and industry literature and describes the use of different types of hydrological models and the uncertainties associated with the hydrological modelling process. This is followed by a discussion into the derivation and application of climate change projections for water resources planning in the UK. The last section reviews recent industry studies and academic research that have considered the impacts of climate change on river flows and the performance of water resource systems. These studies include details of which sources of uncertainty were considered and the approaches used to interpret and present the results from large modelling ensembles. Chapter 2 concludes with a summary of the key research findings and knowledge gaps that have guided the remainder of this thesis.

Chapter 3 introduces the case study catchments along with a description of the hydrological modelling framework, the two conceptual hydrological model structures and the two methods used for estimating potential evapotranspiration (PET) examined in this thesis. The historical climate and flow records for the Thames catchment are also reviewed in Chapter 3.

Chapter 4 describes the development of a hydrological model ensemble for generating flows across the Thames catchment. This development considers two hydrological model structures, along with hydrological model parameter uncertainty, and two methods for estimating potential evapotranspiration (PET) under a Generalised Likelihood Uncertainty Estimation (GLUE) methodology (Beven & Binley, 1992). Prior to its application in later chapters to assess climate change impacts and the performance of London's water resources system, the ensemble's skill (the accuracy and/or degree of association of a model prediction to an observation) in reproducing a range of water resources relevant hydrological characteristics across multiple locations in the Thames catchment is then examined in the remainder of Chapter 4.

Chapter 5 takes forward the hydrological model ensemble and examines the longer term impacts on water resource relevant river flow characteristics using the two climate change products currently recommended for use in water resources planning in England and Wales (Environment Agency, 2012; 2017). These products are the UKCP09 (Murphy et al., 2009b) 'probabilistic' projections (driven by a Medium Emission scenario – see Sexton, 2010) and the eleven-member Future Flows transient climate projections (Prudhomme et al., 2012). This examination includes a comparison between the Future Flows ensemble and the historical record over the 1961 to 2008 period. This is followed by an assessment of the impacts of climate change on river flows across the Thames basin for the 2030s, 2050s and 2080s, and an evaluation of the relative significance of different sources of uncertainty (hydrological model parameter uncertainty, climate model uncertainty and the method used to estimate PET) at each time-horizon.

Chapter 6 extends the findings from chapters 4 and 5 to consider the performance of the water resources system of London using a water resources system model. An approach to sampling the uncertainties across multiple sources and multiple futures is described to enable a pragmatic number of simulations to be identified to estimate the relative impacts of each uncertainty source on the future water supply reliability. These impacts in turn are compared to the range of future demand-side pressures that are projected for London.

Finally, Chapter 7 provides a summary of the research presented in earlier chapters and discusses a number of conclusions of relevance to the UK water resources planning process. Alongside

these, a range of limitations, both with regards to current industry practice and also directly related to the methodological approach adopted for this thesis, are reported. These are supported by suggestions for future research and the additional evidence needed by the industry to enable improved water resources planning in the future.



## Chapter 2 Literature review

The first part of this literature review describes the use of hydrological models for estimating river flows and introduces the different types of models available and the uncertainties associated with the hydrological modelling process. The second section covers the derivation and use of climate change projections, the key uncertainties to be considered in their application, and the availability of climate projections for use by water resource planners in the UK. The last section reviews recent industry studies and academic research that have explored the estimation of climate change impacts on river flows and public water supplies in the UK, including details of which sources of uncertainty were considered and the approaches used to interpret and present results from large modelling ensembles.

### 2.1 Hydrological modelling

A large number of hydrological models have been developed to both improve scientific understanding of hydrological processes and to answer practical questions in relation to flood risk, water resources (including droughts) and water quality. These models can be classified in terms of their underlying approaches to representing real-world hydrological processes and discretisation of the catchment (and climatological inputs).

The performance of a hydrological model is influenced by the choice of model structure (the underlying equations, concepts and numerical schemes), the approach to discretising the catchment and the availability and accuracy of observed datasets (and the choice of performance measures) used to condition/calibrate the model parameters (e.g. Beven, 2001; Clark et al., 2008; McMillan et al., 2010; Krueger et al., 2010; Beven & Smith, 2014). These different sources of uncertainty are complex and involve a lack of knowledge as well as random natural variability (Beven & Smith, 2014) - epistemic as well as aleatory errors (e.g. the non-stationarity of rating curves in measuring river flows to which model outputs will be compared - McMillan et al., 2010; Westerberg et al., 2011b).

The next section describes the different type of models and the sources of uncertainty that should be considered in their application to water resources planning.

#### 2.1.1 *Typology of hydrological models*

Hydrological models can be distinguished by whether they adopt physically-based equations or conceptual representations of the flow processes, and also by how spatial variability is reflected in the model structure.



Lumped, semi-distributed and distributed conceptual models, such as PDM (Moore & Clarke, 1981; Moore, 2007), ARNO (Todini, 1996), TOPMODEL (Beven & Kirkby, 1979; Beven et al., 1984), dynamic TOPMODEL (Beven & Freer, 2001a); CLASSIC (Crooks & Naden, 2007), Grid-to-Grid model (Bell et al., 2007a) and the Thames Catchment Model (TCM or CatchMod; Wilby et al., 1994) represent the hydrological response of a system as conceptual stores in some form without necessarily directly linking model parameters to any “real world” physical process-based equations.

Within the UK water resources planning industry water companies typically use a single model structure as part of their resource assessments. The CatchMod and PDM model structures are extensively used by a number of water companies, as well as in academic research, including a number of studies focused on the Thames Catchment which is the case study used in this thesis (e.g. UKWIR, 2007; Christerson et al., 2012; Matrosov et al., 2013; Thames Water, 2014; Borgomeo et al., 2014; Charlton & Arnell, 2014; Walsh et al., 2015).

In contrast, distributed physically-based models such as SHETRAN (Beven et al., 1987; Abbott et al., 1986), SWAT (Arnold et al., 1998), TOPAKI (Liu et al., 2005) and IHDM (Beven et al., 1987) incorporate systems of non-linear partial differential equations (based on the conservation of mass, momentum and energy) considered to describe all the physical surface and sub-surface flow processes taking place.

In principle, distributed models allow the spatial heterogeneity of inputs, boundary conditions and hydrological processes to be explicitly considered. It is argued (e.g. Refsgaard et al. 1996; Uhlenbrook et al. 2004; Velázquez et al., 2013; McIntyre et al., 2013) that the assessment of land-use and climate change impacts requires the use of such physically-based models to capture the physical changes in the processes that may occur. However, whilst physically-based models may be considered to provide a scientific basis for estimating parameter values (or at least a range between which a value would be expected to reside), few studies (but refer to Bathurst et al., 2004 for an example of a ‘blind’ [without ‘sight’ of measured response data] validation study) appear to have demonstrated the ability to define, a priori, model parameters for any type of model including those which are considered physically based (e.g. Beven, 2001). Bloschl & Montanari (2010) also argue that we cannot hope to reduce uncertainty by including more detail into the models (as in the case of physical, process-based models). The mismatch between model complexity and the availability and quality of data, along with the difficulty in scaling-up physical processes to the necessary coarser computational resolution, means that distributed models, as

with conceptual models, require model “conditioning” (estimation of model parameters) through calibration to observed data (e.g. Beven, 1999).

Another disadvantage of using distributed, physically based models is the increased time required to set-up, calibrate and run the models. This is of particular importance when adopting ensemble and/or risk-based modelling approaches to dealing with uncertainty where very large number of simulations are necessary (e.g. Matrosov et al. 2013; Borgomeo et al. 2014; Walsh et al. 2015).

A further model category is that of metric (or Data-Based Mechanistic, DBM) models (Young & Beven, 2004; Young, 2001) which essentially use the available data (e.g. river flows) to suggest an appropriate model structure without attempting to understand the hydrological processes taking place or spatial heterogeneity across the catchment. In contrast to conceptual and physically based models which effectively represent the modeller’s hypotheses of hydrological behaviour, in metric models the “theory of behaviour” is inferred from observations of the system under study without undue prejudice based on prior hypotheses (Young, 2013). Further details of DBM modelling, how it contrasts and can be used as in conjunction with simulation models (e.g. as an emulator of more complex physically-based models for use in flood forecasting models) is provided by Young & Ratto (2009).

This thesis explores the influence of different sources of uncertainty in the hydrological modelling process using commonly-used conceptual hydrological model structures as part of an ensemble based methodology. Therefore, an overview of conceptual models is presented in the following section.

### *2.1.2 The components of commonly used conceptual hydrological models*

Typically, conceptual hydrological models simplify the hydrological processes taking place into two conceptual phases: run-off generation and run-off routing (Beven, 2001). Run-off generation represents how much of the precipitation becomes part of the flow hydrograph and run-off routing takes into account the distribution of that run-off in time, to form the shape of the hydrograph (Beven & Freer 2001a).

For run-off generation, some models (e.g. SWAT, Arnold et al., 1998; HYSIM, Manley, 2006) adopt infiltration excess mechanisms based on soil infiltration equations such as the Horton (1941) or Green & Ampt (1911; Silburn and Connolly, 1995) equations or empirical relationships such as the Curve Number approach developed in 1954 (Boughton, 1989; Sahu et al., 2010). Alternative approaches to run-off generation used in other models are based on a saturation excess process (e.g. PDM, Moore, 2007; TOPMODEL, Beven & Kirkby, 1979; Beven et al., 1984 - noting that there

are other variants of TOPMODEL which adopt soil infiltration equations), precipitation loss functions (e.g. IHACRES, Jakeman et al., 1990) or water balance models (e.g. CatchMod, Wilby et al., 1994).

It is generally considered (e.g. Todini, 1996; Beven, 2001) that the complexities and nonlinearities of modelling the runoff generation processes are much greater than for modelling the routing processes, with the proportion of run-off for the same event in the same location potentially significantly different due to antecedent conditions. Wagener et al. (2014) suggests that a simple linear routing component, which splits the flow into quick and slow routing components (effectively acting as parallel reservoirs), is usually sufficient for continuous modelling on a daily time-scale.

To allow some of the spatial heterogeneity in climate and hydrogeological characteristics to be captured whilst maintaining relatively rapid simulation times, some conceptual models are resolved on a semi-distributed basis by splitting the catchment into Hydrological Response Units (HRUs) or sub-catchments (e.g. SWAT, HYSIM, HBV and dynamic TOPMODEL) or by discretisation to a coarse grid (e.g. CLASSIC, Crooks & Naden, 2007; the Grid model, Bell & Moore, 1998; Grid-to-Grid, Bell et al., 2007a; WetSpa, Bahremand et al., 2007). These approaches to representing spatial heterogeneity allow different model parameter sets to be adopted in different areas depending upon the dominant flow characteristics in each area. Some of these models (e.g. Grid-to-Grid) extend their representation of spatial heterogeneity further by adopting a variable distribution of soil moisture within each grid cell/HRU and/or by allowing water exchange between cells or HRUs.

Appendix A presents an inventory of several commonly used conceptual hydrological models along with a brief description of their approaches to run-off generation and routing. Further relevant research, in the context of climate change impact assessments, is presented in Section 2.3.

#### *2.1.3 Hydrological uncertainty and model parameterisation*

Hydrological model parameters control or represent the physical processes within a conceptual hydrological model structure in both space and time. There are many different sources of uncertainty, all of which will affect the calibration of model parameters. Since parameter identification will interact with all sources of uncertainty (including scale and commensurability effects), one would expect that values of parameters resulting from calibration would be effective

values that might be different from prior estimations of values based on direct observations or other sources (Beven & Smith, 2014).

The values which parameters adopt are typically derived through training a hydrological model using observed meteorological data to reproduce an observed river gauge record (Boyle et al., 2000). Objective functions such as the volume-error measure (also known as PBIAS - see Gupta et al., 1999), Root Mean Square Error (RMSE, Wagener et al., 2001) and the Nash and Sutcliffe efficiency measure (Nash & Sutcliffe, 1970) have traditionally been used to quantify the degree of agreement between the modelled and observed datasets with model parameters adjusted, either manually or automatically, until an optimal performance is achieved. Other studies (e.g. Blazkova & Beven 2009; Manning et al. 2009; Westerberg et al. 2011a) have evaluated model performance through the “goodness of fit” of specific hydrological signatures (e.g. Sivapalan, 2005; Euser et al., 2013; Westerberg & McMillan, 2015; Broderick et al., 2016), such as flow duration curves, obtained from hydrological data to focus model calibration towards reproducing runoff processes relevant to a particular study.

Euser et al. (2013) highlighted that it is increasingly acknowledged that model evaluation based on single objective optimisation is insufficient to appropriately identify dominant processes and that the use of a multi-objective optimisation potentially offers more insight into the processes underlying the observed catchment response. Adopting multi-objective optimisation, based around different hydrological signatures defined to reflect specific aspects of the system behaviour, potentially offers a better approach for understanding and visualising the strengths and weaknesses of models than using a single aggregate performance measure. Broderick et al. (2016) recommended adopting multiple performance criteria that are pertinent to the study objectives when assessing the transferability of model parameters between contrasting climates.

Dawson et al. (2005) also demonstrated the importance of not relying on individual measures of performance to evaluate data series. Through a comparison of the relative performance of a range of different objective measures across four hypothetical models each exhibiting a different type of error (e.g. timing, reproduction of high or low flows), Dawson et al. (2005) showed that no one model was consistently considered the “best” in terms of the numerous evaluation metrics, although some models appeared to be “better” than others, and that various trade-offs existed.

The choice of the objective function and approach to evaluating model performance can potentially have a significant impact on the perceived performance of a model (e.g. Beven & Freer, 2001b; Cloke et al., 2010). Optimisation or “conditioning” towards a specific function can,

in effect, potentially direct parameter adjustment towards models that reflect specific characteristics of the observed flow record at the expense of other characteristics (e.g. prioritise matching high flows over low flows – see Freer et al., 1996). Consequently, the choice of objective function(s) as well as the choice of hydrological model structure(s) needs to be appropriate to the aims of the research.

#### 2.1.3.1 Equifinality

Beven (2001) suggested that the concept of an optimum parameter set (and model structure) may not be valid with different parameter sets and structures performing to a similar level for a given catchment. As a result of this, Beven (2001) introduced the concept of “equifinality” in hydrology which accepts that there may be many equally valid representations of the hydrological processes and associated parameters in terms of their ability to produce acceptable simulations of the available data.

A number of simple approaches have been suggested to addressing the issue of hydrological modelling uncertainty and equifinality. These include calibration of multiple models to multiple jack-knifed (i.e. systematically removing a single year at a time) historical records in the context of flood risk (e.g. Kay et al., 2009) and varying the values of parameters from a traditionally calibrated model by a fixed percentage (e.g.  $\pm 10\%$ ) for water resources research (e.g. Charlton & Arnell, 2014). These types of approaches effectively explore the sensitivity of parameter values within the calibrated model but do not explore the full feasible parameter space – and it has been shown that sets of similarly performing models can be located in very different parts of the parameter space (e.g. Uhlenbrook et al., 1999; Beven & Freer, 2001b).

This concept of equifinality underpins the Generalised Likelihood Uncertainty Estimation (GLUE) methodology (Beven & Binley, 1992) which is an extension of the regionalised sensitivity analysis of Spear & Hornberger (1980). The GLUE methodology is based on an informal Bayesian method which assumes a prior distribution (often uniform where there is no specific expert knowledge) from which multiple parameter sets are sampled and evaluated (Beven & Freer, 2001b). Given the uncertainties associated with the hydrological model structure, model parameterisation, and historical data, the GLUE approach recognizes that multiple models or model parameter sets may be equally good conceptualisations of the catchment being modelled. Therefore, rather than taking forward a single, considered optimal model to an impact assessment, the GLUE approach leads to what is considered a behavioural model ensemble being used, with each ensemble member associated with a conditional score considered to reflect the relative degree of confidence associated with it.

The original GLUE methodology, which forms the basis of the approach to considering uncertainty adopted in this thesis, has five main steps (Beven & Binley, 1992; Murphy et al., 2006):

- The definition of performance measure(s), chosen on the basis of an objective function(s), to evaluate model performance.
- The definition of a prior distribution for each parameter (typically a uniform distribution for each parameter is assumed at the outset).
- A large number of parameter sets are sampled, commonly using a Monte Carlo Random Sampling (MCRS) approach, from the prior distributions.
- Each parameter set is classified as behavioural (retained) or non-behavioural (rejected) through assessing whether it performs above or below a pre-defined threshold of performance.
- Predictive model runs from each of the retained parameter sets which are combined to generate the conditional distribution of outputs.

As reported by Beven & Binley (2013), the GLUE methodology has been used in hundreds of studies, including research into water resources and climate change impacts (e.g. Cloke et al., 2010; Bastola et al., 2011; Matrosov et al., 2013; Borgomeo et al., 2014; Broderick et al., 2016) as well as flood risk and climate change (e.g. Cameron, 2006; Cloke et al., 2010; Smith et al., 2014).

Whilst the GLUE approach enables uncertainty estimates around model predictions to be quantified, it has been the subject of significant commentary (e.g. Stedinger et al., 2008; Clark et al., 2012; Beven & Binley, 2013) and it is recognised that there continues to be significant academic debate around its application (e.g. Beven & Binley, 2013). The discussion centres on whether it constitutes a rigorous approach to quantifying uncertainty, with the level of debate evidenced by recent publications such as Clark et al. (2012) which consider that GLUE constitutes more of a sensitivity analysis framework (see Pianosi and Wagener, 2016 for discussion on sensitivity analysis frameworks) rather than providing the basis for rigorous quantified uncertainty analysis.

Disagreement primarily relates to the choice of performance measure and acceptability thresholds which must be chosen to reflect model errors in relation to the statistical distribution of the data. As a consequence, other hydrologists (e.g. Clark et al., 2012; Kavetski & Fenicia, 2011; Euser et al., 2013) have promoted alternative approaches such as those based on the use of formal Bayesian statistics. However, as noted by Beven & Binley (2013) formal likelihoods can be used in GLUE if the strong assumptions can be justified (Beven et al., 2008). A key reason for

choosing not to use a formal statistical framework is that real applications may involve significant errors that result from a lack of knowledge (epistemic uncertainties) rather than simple random (aleatory) variability (Beven & Binley, 2013; Beven & Smith, 2014). The reason this is important is that, in general, it is not possible to represent epistemic uncertainties by a formal statistical model with identifiable parameters because epistemic uncertainties will generally result in arbitrarily nonstationary error characteristics (Beven & Smith, 2014). The consequence of treating errors as aleatory when they are significantly epistemic is that the real information content of the calibration data may be overestimated. This may in turn lead to over-conditioning of the model parameters, identifying the wrong model(s) and making incorrect predictions (Beven et al., 2008).

Building upon the GLUE-based philosophy and specific studies that explored different approaches for dealing with observational uncertainty in model evaluation (e.g. Page et al., 2003; Freer et al., 2004), Beven (2006) introduced the concept of “limit of acceptability”. This concept is based around a time-step based performance measure that includes “effective observation errors” that should reflect the different sources of uncertainty in the modelling process for the purpose of model rejection and weighting. These “effective observation errors” include those errors associated with the input data used to drive the models and those associated with the measurements used to evaluate model performance. The “limit of acceptability” approach has been used in several studies (e.g. Liu et al. 2009; Krueger et al. 2010; Coxon et al., 2014) and allows varying limits to be set for individual observations against which to assess model performance. However, in the absence of extensive data, these limits can be difficult to estimate with not all sources of uncertainty being quantifiable. This lack of evidence may inevitably lead to limits greater than the quantifiable uncertainties being specified to avoid the rejection of all model hypotheses (or acceptance that no models will remain within the limits at all time-steps and evaluating models based on their relative score given such limits).

Some studies have combined GLUE and methods such as Bayesian Model Averaging (e.g. Bastola et al., 2011; Broderick et al., 2016). As part of a study to understand hydrological model prediction under contrasting climates, Broderick et al. (2016) applied six alternative hydrological model structures to 37 catchments. For each model structure a GLUE approach was used to consider parameter uncertainty. The median prediction for each was then taken forward as the prediction from that model structure and, building on the research by Bastola et al. (2011), combined using four alternative model averaging techniques (including Bayesian Model Averaging) to examine inter-period transferability. Broderick et al. (2016) demonstrated that a multi-model structure ensemble approach offered improved transferability beyond the reliance on individual model

structures. The authors also highlighted that the performance of model averaging techniques varied in their resource requirements and demonstrated that these model averaging techniques out-performed any of the single model structures used.

The approach to uncertainty, based around a Generalised Likelihood Uncertainty Estimation (GLUE) methodology, adopted in this thesis is described further in Section 4.3.1.2.

#### *2.1.4 The influence of alternative methods for estimating potential evapotranspiration (PET) on model prediction uncertainties*

Previous research (e.g. Kingston et al., 2009; Prudhomme & Williamson, 2013; Seiller & Anctil, 2016) has highlighted that different methods of estimating potential evapotranspiration (PET) can vary significantly. This source of uncertainty may be of particular concern for permeable catchments such as the Thames and may be particularly significant at low flows which is of particular interest to water resources planners (Charlton & Arnell, 2014).

When considering climate change, Kay & Davies (2008) reported that temperature based and empirical approaches led to marked differences in future changes in estimated PET by the end of the 21<sup>st</sup> century using a small ensemble of projections of future climate. Similarly, Ekström et al. (2007) showed large differences between temperature driven and the predominantly physically-based Penman-Monteith (Allen et al., 1998; Droogers & Allen, 2002) formulations in the north-west of England. Finally, Prudhomme & Williamson (2013) compared estimated changes in PET, using 12 different methods including both temperature and predominantly physically-based methods, for a single transient (bias-corrected for temperature) climate projection. Prudhomme & Williamson (2013) demonstrated large differences due to the choice of PET method, with such differences varying seasonally and geographically across the UK.

Kingston et al. (2009) raised the question as to whether it is better to use the more reliable estimates of PET calculated using predominantly physically based equations with potentially uncertain data inputs or empirical methods that only require more reliable data such as temperature.

The influence of the method of estimating PET on hydrological model performance, particularly under climate change, has received limited research attention to date (e.g. Prudhomme & Williamson, 2013) but is likely to be significant (e.g. Haxton & Young, 2012; Seiller & Anctil, 2016). In previous research some studies have made use of predominantly physically-based PET methods such as the Penman-Monteith method (e.g. Wilby & Harris, 2006; Cloke et al., 2010; Bell et al., 2012; Charlton & Arnell, 2013), whilst others have used temperature based PET methods (e.g.



Christierson et al., 2012; Kay & Jones, 2012; Cloke et al., 2013). However, none of these studies have compared multiple formulations and considered the potential implication on the projected impacts of interest to water resources planners. Further, few previous studies assessing the hydrological impacts of climate change have considered the feedbacks between climate change, vegetation and PET. A notable exception is the research by Bell et al. (2011) which allowed surface resistance to vary, simulating the closure of plant stomata under higher atmospheric CO<sub>2</sub> concentrations, to demonstrate much lower increases in PET as a result of a climate change.

It is hypothesised (Manning et al. 2009; Charlton & Arnell 2014) that such estimates of PET are likely to be particularly important when considering water resources in catchments, such as the Thames, where the balance between precipitation and PET is critical to the availability of water. Consequently, this knowledge gap is examined throughout this thesis, with details of the approaches adopted for estimating PET described in Chapter 3.

## 2.2 Climate change projections for water resources planning in the UK

Jenkins et al. (2009) summarised that the evidence of hydrological trends are consistent with future climate change scenarios, with recent observations of rising sea levels and warmer conditions in the UK over the last few decades. Kendon et al. (2016) also reported that there has been an increase in temperature from the 1970s to the 2000s with the most recent decade (2006–2015) being, on average, 0.9 °C warmer than the 1961–1990 average.

In terms of precipitation, Jenkins et al. (2009) reported that whilst annual mean precipitation over England and Wales has not changed significantly since records began in 1766, precipitation amounts appear to have decreased in summer and increased in winter. Kendon et al. (2016) however, did report a slight increasing trend from the 1970s onwards for England and Wales, with the most recent decade (2006–2015) on average 7% (England) and 5% (Wales) wetter than 1961–1990.

From a water resources perspective, Marsh et al. (2014) reported that for much of the 19th century, summer rainfall exceeded that for the winter and clusters of dry winters were common. By comparison winter rainfall for England and Wales over the last 30 years (to mid-2000s) has been around 10% above the previous long-term average, and summer rainfall around 10% below, potentially leading to increased annual runoff and groundwater recharge (due to greater proportion of the rainfall occurring when evaporative demands are lower).

Previous UKWIR-sponsored research (2007) reviewed trends in 47 river flow records and 10 groundwater observation wells and found a small number of short to medium-term upward trends in winter and autumn runoff. However, the study reported that these changes could not be attributed to climate change over and above natural variability and the study suggested that it will take some time before a clear (and statistically robust) picture emerges of observed changes in river flows. This conclusion was further emphasised by Watts et al. (2015) which reported that there is currently insufficient evidence to suggest a link between anthropogenic climate change and changes in precipitation, river flows and groundwater levels in the UK.

Hannaford (2015) reviewed the available research evidence for long-term climate-driven changes in UK river flows, considering both high and low flows, and found that changes can be detected in river flow regimes, some of which agree with climate change projections, while others are in apparent contradiction. For low flows in particular, Hannaford (2015) concluded that, overall, there seems to be little evidence of any strong decrease in low flows since the 1960s. Hannaford (2015) also confirmed the commentary by Watts et al. (2015) that observed changes generally cannot be attributed to climate change, largely due to the fact that river flow records are limited in length and the identification of short-term trends is confounded by natural variability.

The next section provides an overview of the methods used to develop climate change projections, the sources of their associated uncertainties and the approaches typically used to apply projections in impact studies. In addition, a brief history of the climate change projections developed for the UK is provided, including a summary of the UKCP09 and Future Flows climate projections referred to in the latest Water Resources Planning Guidance (Environment Agency, 2012; 2017).

#### *2.2.1 Global Climate Models (GCMs) / General Circulation Models*

General Circulation Models are used to represent the physical processes in the Earth's atmosphere or ocean and form the basis of Global Climate Models (GCMs) which are numerical coupled models that represent the various earth systems including the atmosphere, oceans, land surface and sea ice. GCMs are considered to offer considerable potential for the study of climate change and variability (e.g. Fowler et al., 2007a) but their projections are subject to several sources of uncertainty (e.g. Collins, 2007) including:

- uncertainty in climate models (e.g. feedback processes, land–atmosphere and ocean–atmosphere coupling);
- uncertain pathways for forcing agents such as greenhouse gas and particulates emissions;

- natural climatic variability.

#### 2.2.1.1 Climate model uncertainty

It is recognised (e.g. Murphy et al., 2009b) that there are a number of limitations with existing climate model structures due to current scientific understanding, computational limitations, and available datasets.

Global Climate Models (GCMs) typically work at a temporal resolution of 30 minutes and a horizontal resolution of 250 km (IPCC, 2013) which is coarser than the scale at which some climate system processes (e.g. convective clouds) occur. As a result, such sub-grid processes need to be parameterised within the GCM large-scale responses with, for example, precipitation assumed to occur at a uniform rate everywhere within a GCM cell. This can lead to an overestimation of rainfall frequencies and an underestimation of intensities compared with reality and it is recognised that climate models typically show greater consensus (i.e. a smaller range of uncertainty) in their projections of temperature changes than for precipitation (e.g. Stainforth et al., 2005; Deser et al., 2012; Murphy et al., 2009b). This is primarily due to the difficulties in modelling the processes and mechanisms used to generate precipitation and reproducing blocking events (see section 2.2.1.2) in the climate models, which is at least, in part, due to their coarser resolution relative to the physical processes involved (e.g. Maraun et al., 2010; Cloke et al., 2013).

A further issue is that the global or regional metrics typically used to condition GCMs may not be appropriate for a local scale impacts study (e.g. Wilby & Harris, 2006), particularly in the case of precipitation, and therefore it is necessary to evaluate how climate sequences of relevance to an impact study are reproduced in comparison to the observed records.

Climate model uncertainty can be sub-divided into structural uncertainty (process representation) and parameter uncertainty (Collins, 2007) with multi-model ensembles (MMEs) used to explore structural uncertainty (Solomon et al., 2007; Murphy et al., 2009b) and perturbed physics ensembles (PPEs) used to explore parameter uncertainty (Murphy et al., 2009b; CPDN, 2016).

These MME and PPE studies have demonstrated that different climate models and parameter sets can produce significantly different projections of future climate for the same emission scenario, with in some cases, simulations disagreeing in the sign of predicted changes in some regions. This is despite these models all being considered to adequately reproduce the historical climate. Consequently, explicitly incorporating these different sources of uncertainty into climate predictions is considered critical and has driven the development of ensemble and probabilistic techniques in climate prediction and impact studies (e.g. Collins, 2007; New et al., 2007a).

#### 2.2.1.2 Natural climate variability

Typically, water resources planning takes into account natural climate variability through the use of long historical records intended to cover a large range of possible climate sequences including periods of drought conditions or extreme storm events. In some studies (e.g. Crooks & Naden, 2007; Prudhomme & Davies, 2009) historical records have been resampled in an attempt to better understand the potential range of natural climate variability (which may not be directly present in the observed records) so that it can be better distinguished from any climate change signal.

In the context of water resources, Jones et al. (2009) reported that one type of extreme that is currently particularly difficult for GCMs to reproduce is spells of similar weather patterns. These events relate to the persistence of specific weather types, often referred to by climatologists as blocking or weather regimes. These events are particularly important with respect to both heatwaves/droughts (e.g. the 1976 summer drought in the UK) and also exceptionally cold winters (e.g. the UK winter of 1962–1963) and climate models have typically underestimated the occurrence of blocking in the Euro-Atlantic sector (Scaife et al., 2011). Sillmann et al. (2016) also highlighted that current climate models have large biases in some regions and may not be able to adequately simulate key dynamical patterns such as atmospheric blocking or other weather regimes, jet stream position and intensity, tropical dynamics and teleconnections, or stratosphere-troposphere connections.

Looking ahead, Anstey et al. (2012) noted that the CMIP5 GCMs that will underpin the next set of climate projections for the UK (UKCP18), continue to underestimate the observed blocking frequency over Europe with the distribution of biases similar to those associated with the CMIP3 models that underpin the UKCP09 projections currently used by water resources planners in the UK (see section 2.2.2.1).

The current water resources planning guidelines (Environment Agency, 2012; 2017) for England and Wales recommends that water companies adopt the 2030s or 2080s time-horizons for climate change (Environment Agency, 2012; 2017). Wilby et al. (2009; 2011) suggested that, over shorter time-horizons (e.g. 2020s compared to 2080s), any climate change signal present in projections may be weak relative to natural climate variability and that such natural internal climate variability will be magnified at the finer spatial scales used for impact studies, potentially increasing the uncertainty in predictions. Consequently considering multiple, future time-horizons may better distinguish the climate change signal from natural climate variability.

#### 2.2.1.3 Downscaling of climate projections for impacts modelling

The low spatial resolutions of GCMs make them difficult to apply directly to the river basin scale applications required for flood risk and water resources planning in the UK. The coarse scale of GCMs means that they do not provide information as to the spatial structure of temperature and precipitation in areas of complex topography and land use distribution. Also their depiction of regional and local atmospheric circulations and representation of processes at high frequency temporal scales are considered insufficient for regional scale impacts modelling (e.g. Christensen et al., 2007).

Fowler et al. (2007a) suggested that, in particular, the reproduction of observed spatial patterns of precipitation and daily precipitation variability by GCMs is not sufficient for hydrological impact modelling directly. The process of resolving the resolution of climate models and regional and local scale processes is termed 'downscaling'. The downscaling of climate information to the local (impact model) scale is typically achieved using the either Regional Climate Models or empirical / statistical downscaling techniques.

#### 2.2.1.4 Dynamical downscaling and the used of Regional Climate models

Dynamical downscaling refers to the use of Regional Climate Models (RCMs) that use large-scale lateral boundary conditions from GCMs to provide higher resolution outputs (e.g. Fowler et al., 2007a; Lafon et al., 2013). These RCMs are typically resolved at the 25 to 50km latitude and longitude scale (e.g. Smith et al., 2014) and parameterise physical atmospheric processes and are thus considered able to realistically simulate regional climate features such as orographic precipitation, extreme climate events and regional scale climate anomalies and non-linear effects (Fowler et al., 2007a; Lafon et al., 2013).

The output from RCMs is coherent both spatially and temporally (Smith et al., 2014), in the sense that they arise from model(s) which produce dynamically and physically consistent simulations of

the passage of a sequence of atmospheric weather systems. This means, for example, that daily data from any number of squares (contiguous or otherwise) can simply be spatially aggregated to form a physically plausible area average over any desired region, including river basins (Murphy et al., 2009b). Dynamical downscaling is also attractive if large numbers of jointly varying outputs are required although further downscaling may be required to provide greater spatial detail including the effects of local scale topography.

RCM simulations typically provide projections of absolute future climate rather than changes relative to the baseline. As with all downscaling approaches, outputs from RCMs will contain biases, due to different types of systematic errors within their formulation and inherited larger scale biases from their driving global simulations (Murphy et al., 2009b). Consequently, bias-correction (e.g. Bell et al., 2007; Lafon et al., 2012; Cloke et al., 2013) is typically undertaken for hydrological studies that use outputs from Global and Regional climate models.

The effect of historical model biases from climate models can be partially removed by differencing a baseline period from the future projections to get a climate change with which to perturb an observed climatology (e.g. Arnell et al., 2003). However, this does not mean that the future projections will be error free due to the uncertainty in modelling, but that the historical model bias will have been removed. It is worth noting that this method only corrects biases in the mean and that biases in other statistical aspects (e.g. variance and extremes) may still remain (Lafon et al., 2013). In addition, the bias from RCMs may lead to implausible results for the present climate from the impacts model, in which case a bias adjustment to the impacts by subtracting the present from the future may be inappropriate (Murphy et al., 2009b).

A number of studies (e.g. Teutschbein & Seibert, 2013; Lafon et al., 2013; Smith et al., 2014) have examined a range of bias-correction approaches to provide evidence as to how their suitability can be assessed and how associated limitations might influence modelling results and conclusions. Lafon et al. (2013) examined four published bias-correction techniques for precipitation generated using a regional climate model (HADRM3.0-PPE-UK) for seven sites in the UK and suggested that a gamma distribution-based quantile mapping approach may offer the best combination of accuracy and robustness. However, Lafon et al. (2013) also highlighted the challenges for all bias-correction techniques in correcting the third and fourth statistical moments (i.e. skewness and kurtosis) and the sensitivity of the bias-correction to the calibration period used, which is of particular interest when considering the extremes. Further, the use of such bias-correction techniques makes the assumption that the physical processes that lead to precipitation will not change into the future and will also be affected by the same uncertainties associated with

the historical records (e.g. Teutschbein & Seibert, 2013). Perhaps, more critically, if the underlying RCMs are not able to provide credible projections of future climate directly, and it is considered necessary to employ significant bias-correction prior to their use in impacts modelling, then there may be a question as to the RCMs' 'fitness-for-purpose'. Transforming an output such as precipitation is effectively 'throwing away the physics' that underpin the RCMs and therefore the reasons why these models might be considered able to provide feasible futures in the first place (e.g. Cloke et al., 2013).

#### 2.2.1.5 Statistical downscaling

A simple statistical downscaling approach involves the use of change factors (perturbation method) derived by comparing the baseline and projected climatology produced by a climate model which are then applied to an observed climatological record.

The use of a change factor approach to impacts modelling is commonly used (e.g. Wilby & Harris, 2006; New et al., 2007a; UKWIR, 2007; Reynard et al., 2010) primarily because of its simplicity and that, in principle, it combines one-step downscaling and bias correction (Willems et al., 2012), assuming that relative changes from climate models are more reliable than absolute values (Ntegeka et al., 2014). However, change factors typically only reflect changes in the mean climate (e.g. Fowler & Kilsby 2007) with future variability considered unaffected and the sequence of weather patterns unchanged. Using such change factors therefore makes a fundamental assumption that changes in weather during extreme events (such as droughts) will be similar to changes in the mean climate which may not necessarily be the case. In addition, the timing and profile of drought sequences are preserved in the climatology and therefore the water resources system to be tested will only be subjected to changes to the intensity of events in the historical record and not to new and previously not encountered weather sequences. A final point is that the climate is considered stationary during each epoch being considered and therefore multiple epochs need to be considered to develop a trajectory of climate change impacts for a given system.

In applying climate model outputs using the change factor approach, the World Meteorological Organisation recommends a 30-year period as a reference climate. A 30-year period is recommended because it is expected to contain enough climate variability to provide a robust estimate of mean climate with the climate change signal considered too small compared to natural variability to introduce bias in the calculation of the average.

Reynard & Prudhomme (2010) highlighted the influence that the choice of baseline period can have on the change factors calculated. For several GCMs, Reynard and Prudhomme (2010) calculated monthly rainfall changes as the difference between a fixed period (2071-2100) of the future GCM run, and each of the 30-year periods resampled from within the control run (1951-2000) to provide a range of potential delta change factors. When the monthly change factors, defined strictly as per the method recommended in the IPCC-Third Assessment Report (i.e. 2071-2100 minus 1961-1990), were compared with this range, it was observed that the IPCC- TAR factors were sometimes outside the 50% band around the median of the range of all factors.

Reynard & Prudhomme (2010) also explored the influence of shorter averaging periods (10 and 20 years) on the calculated monthly change factors and indicated that shorter averaging periods demonstrated larger variability. This leads to the report concluding that relying on one single definition of the factors may present a chance that an important source of uncertainty in climate change projections resulting from natural climate variability is ignored.

In another study, Lopez et al. (2009) adopted a gamma transform approach (Wood et al., 2004), also known as quantile to quantile mapping, to correct for GCM bias with a gamma distribution fitted to both the observed and GCM outputs for a water resources study. The gamma transform approach is conservative in that it is impossible to obtain values outside the range of the observation data which will particularly affect the distribution of the extreme monthly precipitation since these will be mapped to the extreme values within the climatology.

In terms of other statistical downscaling approaches, Diaz-Nieto & Wilby (2005) compared the relative merits of a change factor approach with a statistical downscaling approach applying climate variable from GCMs using transfer functions to estimate point-scale meteorological series. This study considered low flows in the River Thames using climate change outputs from the UKCIP02 projections (see section 2.2.2). A key assumption in adopting the transfer functions is that it must be assumed that the empirical relationships linking large-scale predictors to local predictand(s) are valid under future climate forcing. The potential benefit is that the approach can provide ensembles of daily climate that evolve in line with the large-scale, transient changes of the driving GCM.

The impact of the different approaches was that changes in flow associated with using the transfer functions were generally more conservative and complex than that arising using change factors. Diaz-Nieto & Wilby (2005) suggested that these departures are explained in terms of the different treatment of multi-decadal natural variability, temporal structuring of daily climate



variables and large-scale forcing of local precipitation and potential evapotranspiration. Diaz-Nieto & Wilby (2005) highlighted striking differences between the approaches in terms of changes in potential evapotranspiration (a Penman method was used), with those generated using change factors significantly higher than derived using the transfer functions. Finally, Diaz-Nieto & Wilby (2005) commented that the use of the statistical transfer functions is more time-consuming (compared to using change factors), particularly where spatially coherent outputs across multiple sites are required, such as that described in Yates et al. (2015) in their application to a water management study in Colorado (refer to 2.3 for further discussion of this research).

Finally, Clark et al. (2016) reported that non-stationarity in statistical downscaling model parameters (and the impact in turn on predictor to predictand relationships) is widely recognised as a key problem and has yet to be seriously characterised or resolved by the community, creating considerable uncertainty in how climate change is portrayed. On this key point, Clark et al. (2016) made reference to the research undertaken by Charles et al. (1999) that highlighted that the validation of a statistical downscaling technique for present day conditions does not necessarily imply legitimacy for changed climate conditions and that statistical downscaling studies that have not attempted to determine the plausibility of their predictions for the changed climate conditions should be viewed with caution.

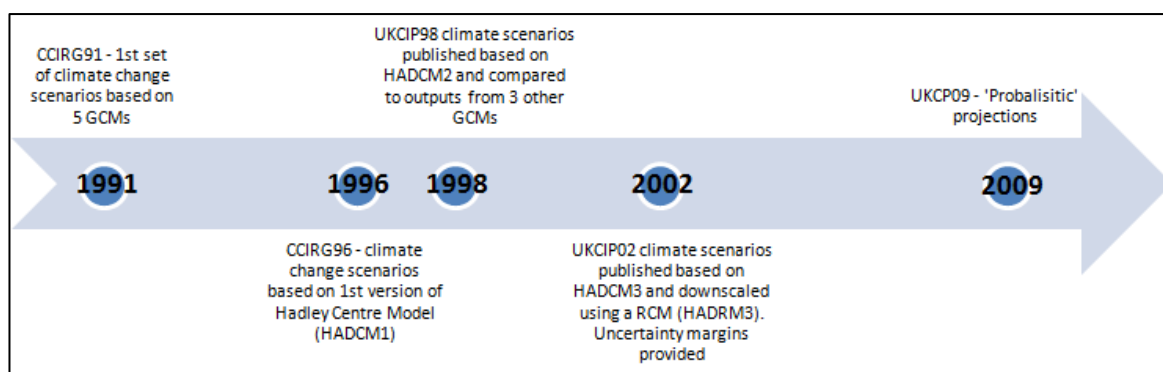
### *2.2.2 History of climate change projections in the UK*

UK Climate Projections (UKCP09) represents the fifth generation of climate scenarios that have been produced for the UK since the publication of the first scenarios (CCIRG91) in 1991. The scenarios have evolved as scientists' knowledge of the climate, available computing power and stakeholder needs have evolved as summarised in Figure 2-1 (based on a similar image available at UKCIP, 2010). For the UK, UKCP09, published in June 2009, was the first attempt to provide probabilistic climate change projections.

#### *2.2.2.1 UKCP09 Probabilistic climate projections*

One of the products provided as part of the UKCP09 projections (Murphy et al., 2009a) is a 10,000-member 'probabilistic' ensemble of seven, stationary, 30 year climate projections, from the 2020s (2010 to 2039) to the 2080s (2070 to 2099) over a 25km grid, for 16 administrative regions, 23 river-basins and 9 marine regions. These projections are provided as monthly, seasonal and annual averages of a range of climate variables including precipitation and temperature over each 30 year period as summarised in Figure 2-2 (adapted from Murphy et al., 2009a). With regards to the calculation of potential evapotranspiration, projected changes in wind speed were not initially available as part of the UKCP09 projections (see Prudhomme &

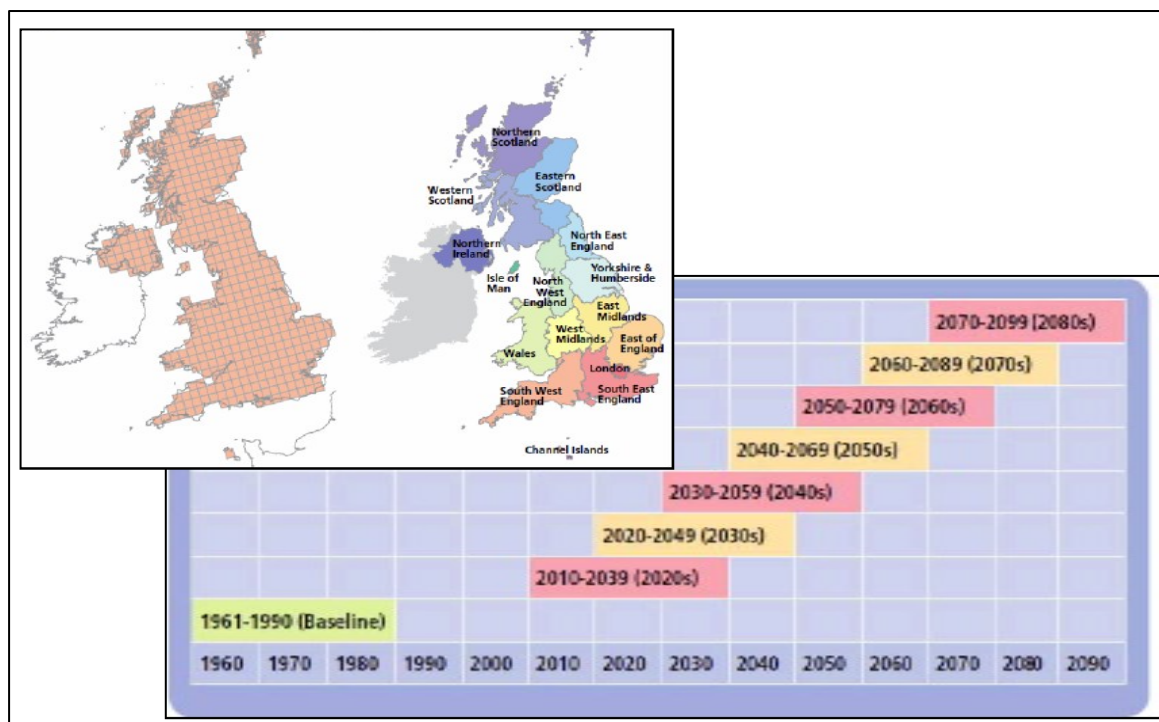
Williamson, 2013, for further details) although they were subsequently published as a separate product which cannot be used in conjunction with other UKCP09 variables (see UKCP09, 2017). It should be noted that the change factors are derived from the change in mean monthly values over the 30 year time-slices relative to a 1961 to 1990 baseline.



**Figure 2-1: Timeline of the publication of climate change scenarios for the UK.**

To comprehensively allow for all sources of potential uncertainty in climate modelling is beyond current resource availability. Consequently, for UKCP09, the Met Office developed an approach to account for the major known sources of uncertainty in future projections by combining results from perturbed variants (Perturbed Physics Ensembles) of the HadCM3 configuration of the Met Office global climate model with projections from alternative international climate models using a statistical emulator (Sexton and Harris, 2015). RCMs were used to provide higher resolution climate projections consistent with the driving global model projection at finer (25 km) spatial scales (Murphy et al., 2009b).

Current approaches to deriving probabilistic projections typically use global, regional or local metrics to weight GCM model outputs. Sexton & Harris (2015) summarised the generation of probability density functions (PDFs) using Bayes' theorem, as used to generate the UKCP09 "probabilistic" projections. The term "probabilistic" has aroused controversy in that such "probabilistic" climate change projections are constructed on the basis of available evidence and judgement and are not estimates of a 'true' distribution (e.g. Stainforth et al., 2007). Stainforth et al. (2007) argued that such a probabilistic approach can be misleading because of the non-stationarity of the climate system, with past observations not having sampled the full state space and GCMs having known limitations with regards to spatial coherence, interannual-variability and blocking/extremes - all of which are of particular interest for water resources studies. In addition, there may be other currently unknown uncertainties that are not included in the GCM formulations.



**Figure 2-2: Overview of UKCP09 climate change projections.**

### 2.2.2.2 Eleven member UKCP09 Regional Climate Model ensemble and the Future Flows projections

A perturbed physics ensemble of eleven Met Office regional climate model (Hadley Centre Regional Climate Model (HadRM3) variants were run (and retained) as part of UKCP09 over a domain significantly larger than the spatial extent required for the final projections. This larger domain was chosen to avoid the risk that relaxation to GCM data at the lateral boundaries will dampen the simulation of finer scale detail over interior regions of interest (Murphy et al., 2009b). Parameter settings in each RCM ensemble member were chosen to be consistent with the settings used in the relevant HadCM3 simulation. This RCM ensemble was run from 1950-2099, driven by the Medium Emission scenario SRES A1B (Murphy et al., 2009b) of future emissions of greenhouse gases and aerosols, and provides climatic projections at 25 km resolution - there are no corresponding RCM projections for other emissions scenarios. This RCM product provides daily, absolute values (not relative climate changes), continuous from 1950 to 2099, which are spatially and temporally coherent.

The Future Flows projections (Prudhomme et al., 2012) derived an eleven-member ensemble of transient future climate projections specifically developed for hydrological and hydrogeological modelling based on these eleven member regional climate model runs. Prudhomme et al. (2012) reported that a statistically-based bias-correction and downscaling procedure (Newton et al., 2012) was implemented nationally over Britain to generate the Future Flows Climate series. A

linear additive transfer function (Leander and Buishand, 2007), based on the 5-km daily temperature time series UKCP09 gridded observation datasets (Perry et al., 2009), was applied to the HadRM3-PPE RCM temperature time series at the 5-km resolution of the observations for each month. The parameters of the transfer function were estimated so that, for each 5-km grid cell and over the period 1962–2000, mean monthly bias-corrected temperature matched the mean monthly observed temperature over the same area.

For precipitation, a detailed review of possible approaches was undertaken as part of the Future Flows project (Newton et al., 2012) with the final choice of transfer function following the parametric quantile-mapping method described by Piani & Haerter (2012), based on a gamma distribution. 25-km average observed daily precipitation, matching the resolution of HadRM3-PPE aggregated from the 1-km daily observed precipitation time series, was used to establish the transfer function for each month. Each 25-km grid of the averaged observational data and most of the statistical properties (and in particular the first two statistical moments of daily rainfall) of the bias-corrected output were considered suitably similar to the observed records over the 1962–1991 historical period (Prudhomme et al., 2012). For estimating PET, the Future Flows project employed the predominantly physically-based FAO-56 Penman Monteith method (Allen et al., 1998; Droogers & Allen, 2002) at a monthly time scale following a detailed assessment of different methods (further details of which can be found in Prudhomme & Williamson, 2013).

In contrast to the use of change factors, such as the UKCP09 projections, the use of the transient time-series such as the Future Flows projections does allow the testing of a system under weather sequences not present in the historical record. However, there is currently no published evidence as to how extreme events of interest to water resources planners, such as droughts, are reproduced by these transient projections (e.g. downscaled and bias-corrected outputs from Regional Climate Models) and therefore how representative they are of the future risks of such events under a changing climate. The climate models that underpin both the UKCP09 probabilistic and Future Flows projections are recognised as having limitations in reproducing blocking patterns that are significant in the formation of extended droughts (e.g. Rocheta et al., 2014; Watts et al., 2015). The need for bias-correction emphasises limitations associated with the underlying RCMs and is itself a further source of uncertainty (e.g. dealing with skewness and kurtosis and the sensitivity to the choice of calibration period) with regards to droughts and water resources planning (Teutschbien & Siebert, 2013; Lafon et al., 2013). A further limitation of the Future Flows product is that it contains only eleven members, all using the same underlying

climate model, and therefore can only reflect a subset of the known uncertainties (Prudhomme et al., 2012).

The literature supporting the publication of the Future Flows projections (Prudhomme et al., 2012) recommends that, in impact studies, each ensemble member is considered equally likely, though there is currently little evidence published where this assumption has been demonstrated as being appropriate. An alternative approach is to weight such projections according to a user defined criterion such as that adopted by Wilby & Harris (2006) who weighted projections based on their performance under current conditions to observations according to what they considered a hydrologically relevant indicator. In their study, Wilby & Harris (2006) considered each projection's bias compared with observations in reproducing downscaled summer (June to August) effective rainfall (precipitation minus PET). One of the challenges of doing this is in defining appropriate criteria to minimise the risk of emphasising projections which may reproduce some outputs well but for the wrong reasons. A more rigorous approach might be to examine the key underlying processes used in the development of the projections but these are not typically available to the impacts modeller. At the very least, where possible, such projections for use in water resources planning should be examined in the context of their proposed application prior to their use.

### 2.3 Assessing the hydrological impacts of climate change for water resources planning

The majority of previous research that has considered water resources under climate change has predominantly focussed on the impacts of uncertainty sources on river flows (e.g. Christiernson et al.; 2012, Prudhomme et al., 2012; Borgomeo et al., 2014; Walsh et al., 2015) and have rarely taken their analysis as far as quantifying the impacts upon specific water resources planning decision-relevant metrics. Where such studies have done this to some degree (e.g. Fowler et al. 2007b; Lopez et al., 2009; Borgomeo et al., 2014; Turner et al., 2014; Walsh et al. 2015), only some of uncertainty sources have been considered, with the respective research scopes typically limited by resource constraints and access to models and data.

The remainder of this section reviews a number of recent industry studies and academic research that have considered the hydrological impacts of climate change that are of particular relevance to this thesis. This review is divided into separate sections looking at industry-focussed studies and academic research, with a summary of the models and methodologies used presented in Table 2-1.

Table 2-1: Selected research studies investigating the hydrological impact of climate change (with an emphasis on UK studies of relevance to the water resources planning industry).

Research reference	Catchments	Projected time horizon	Quantified uncertainty	Emission scenarios	GCM	RCM	Downscaling	PET equation	Hydrological Modelling				Water Resource Systems Modelling
									Resolution of model and input climate data	Time-step	Model structures	Model parameter sets	
Borgomeo et al., 2014	1 – Thames to Kingston	2050s	GLUE (Beven & Binley, 1992)	1 – SRES A1B	Based on UKCP09 probability distribution			Not reported	Lumped	Daily	CatchMod	53 – weights based on performance assessed against volume error (+/- 10%)	Yes – London WRZ
Charlton & Arnell, 2014	6 UK catchments	2020s, 2050s and 2080s	No	3 SRES (B1, A1B, A1F1)	Based on UKCP09 probability distribution			Penman-Monteith (with wind speed unchanged under future scenarios)	Lumped	Daily	Cat-PDM	1 – based on performance at low, average and high flows	No
Christierson et al., 2012	70 UK catchments	2020s	GLUE	1 – SRES A1B	Based on UKCP09 probability distribution			Oudin (Oudin et al., 2005)	CatchMod – 3 HRUs.  PDM lumped	Daily flow sequences	2 - CatchMod / PDM	10,000 sets - NSE(Nash and Sutcliffe, 1970) on logarithm of flows with models > 0.5 retained	No
Cloke et al., 2010	1 – Medway, England	2080s	GLUE (Beven & Binley, 1992)	1 – SRES A1B	1 – HADCM3	1 RCM (HADRM3) – 10 variants	Direct RCM data (bias-correction explored but not used in final analysis)	Not reported	Lumped catchment  Catchment average climate data	Daily	1 - CatchMod	Likelihood weighted flow distributions - behavioural models selection and weighting based on Nash-Sutcliffe (1970) score > 0.6	No
Fowler et al., 2007b	North-West England IRZ	2070 – 2100 - pattern scaling used for other time- slices	No	1 SRES (A2)	1 GCM (HADCM3)	1 RCM (HADRM3H) – pattern scaling used for other emission scenarios	Bias-corrected RCM output used directly	Blaney-Criddle (1950)	Spatially uniform climatic data  Basin discretisation not reported	Daily flow sequences	1 – ARNO (simplified)	1 set - optimised to NSE (Nash and Sutcliffe, 1970) using SCE (Duan & Qingqyun ,1992)	Yes – Integrated WRZ
Fowler et al., 2008	1 - Eden, catchment	2050s	Yes – Bayesian	1 SRES (A2)	2 (HADAM3H/ HADAM4P/ HADCM3 and ECHAM4/O PYC3)	13 RCMS (PRUDENCE ensemble)	Weather Generator (EARWIG) using change factors (monthly) to downscale from RCMs - 1000 30 year sequences for each RCM generated	Not reported	Areal average climatic data  Basin discretisation not reported	Daily flow sequences	1 – ARNO (simplified)	1 set - optimised to NSE (Nash and Sutcliffe, 1970) and water balance using SCE (Duan & Qingqyun, 1992)	No
Lopez et al., 2009	South-west England	2000 - 2080	Unweighted model assessments	1 SRES (A1B)	1 PPE for GCM (HADCM3L) from CPDN and GCM ensemble (CMIP3)	None	Gamma transform Quantile to quantile mapping) on a monthly basis	Penman (1948) (simplified - with wind speed unchanged under future scenarios)	1 HRU  Areal average climatic data	Daily flow sequences	1 – CatchMod	1 set (parameters provided by Environment Agency)	Yes – Wimbleball WRZ

Research reference	Catchments	Projected time horizon	Quantified uncertainty	Emission scenarios	GCM	RCM	Downscaling	PET equation	Hydrological Modelling				Water Resource Systems Modelling
									Resolution of model and input climate data	Time-step	Model structures	Model parameter sets	
Manning et al., 2009	1 – River Thames to Kingston as set-up under Wilby & Harris (2006)	2020s/2050s/2080s	Yes (Bayesian – Regional observation climate metrics & ensemble consensus on future temp. mean)	4 SRES (A1F1,A2, B1, B2)	2 –HADCM3 and OPYC/ECHA M4	14 RCMs (HADRM3H + 13 PRUDENCE RCMs)	Synthetic time series derived from RCM outputs (using Weather Generator)	FAO-modified Penman (1994) or MORECS Penman-Monteith (Hough & Jones, 1997)	3 HRUs  Areal representative climatic data	Daily flow sequences	1 – CatchMod	12 (pareto-optimal) sets based on the NS (Nash and Sutcliffe, 1970) and absolute flow	No
Matrosov et al., 2013	1 – Thames to Kingston	2030s	GLUE (Beven & Binley, 1992)	3 SRES (B1, A1B, A1F1)	Based on UKCP09 probability distribution			Oudin (Oudin et al., 2005)	CatchMod – 3 HRUs.  PDM lumped	Daily flow sequences	2 – CatchMod / PDM	Not reported but based on ensemble from Christierson et al., (2012) with NSE >0.5	Yes – London WRZ
New et al., 2007a	1 – Thames to Kingston	Future equilibrium at 2 X CO <sub>2</sub>	Yes	1 (2 X CO <sub>2</sub> )	1 – HADSM3 2700 parameter sets	None	Perturbed baseline with monthly change factors	Penman (1948)	3 HRUs	Daily flow sequences	1 - CatchMod	100 sets (as defined in Wilby & Harris, 2006)	No
Prudhomme et al., 2005	13 case study catchments across UK	2020s (some 2080s where 2020s not possible)	Unweighted	2 SRES (A2 and B2)	3 - HadCM3, CCGCM, CSIRO-Mk2)	HADRM3H	3 - Statistical (SDSM), dynamical (HadRM3, only for 2080s) and delta method (UKCIP02)	Penman-Monteith(Allen et al., 1998)	Single spatial unit  Areal representative climatic data	Daily flow sequences	2 versions of PDM	Monte-Carlo based on several objective functions – on average 110 per catchment	Yes – 2 conceptual systems
Prudhomme & Davies, 2009	4 test catchments across UK	2080s	No	2 SRES (A2 and B2)	3 - HadCM3, CCGCM2, CSIRO)	HADRM3	3 – dynamical (HADRM3) and statistical (SDSM) from GCMs. Also delta method applied to HADRM3 outputs	Penman Monteith (Monteith, 1965)	Areal representative climatic data	Daily flow sequences	2 versions of PDM	Nash and Sutcliffe, 1970	No
Prudhomme et al., 2009	282 UK catchments	Transient from 1950 to 2099	No	1 SRES (A1B)	HADCM3	HADRM3	Downscaled to 1 or 5km grid with bias-correction	Penman-Monteith (Allen, 1998)	PDM – Lumped, CERF/CLASSIC Gridded	Daily	PDM, CLASSIC, CERF	1 – calibrated (NSE) for PDM, Regionalised (CERF) and Part-regionalised, part-calibrated (CLASSIC)	No
Turner et al., 2014	Melbourne, Australia	2025 to 2055	No	6 - A1B, A2, B1, B2, A1F1, and AT (Nakicenovic & Swart 2000)	23 (see Turner et al., 2014)	None	Patterns in Mean Annual Temperature and Precipitation changes	Not used	No hydrological modelling	N/A	N/A	N/A	Yes
UKWIR, 2007	70 UK catchments	2020s	GLUE (Beven & Binley, 1992)	1 – SRES (A2)	6 - HadCM3, CGCM2, CSIRO-mk2, GFDL-R30,ECHAM 4/OPYC3,	None	Disaggregation of bias-corrected (quantile to quantile mapping) of	(Oudin et al., 2005)	CatchMod – 3 HRUs.	Daily flow sequences	2 - CatchMod / PDM	10,000 sets - NS (Nash and Sutcliffe, 1970) on logarithm of flows with models > 0.5 retained	No

Research reference	Catchments	Projected time horizon	Quantified uncertainty	Emission scenarios	GCM	RCM	Downscaling	PET equation	Hydrological Modelling				Water Resource Systems Modelling
									Resolution of model and input climate data	Time-step	Model structures	Model parameter sets	
					CCSR/NIES		GCM outputs using monthly spatial anomalies		PDM lumped				
Walsh et al., 2015	1 – Thames to Kingston	2020s / 2050s	No	1 – SRES A1B	Based on UKCP09 probability distribution and application of an extended Weather Generator			Not reported	3 HRUs with climate inputs catchment average	Daily flow sequences	1 - CatchMod	1 for each inflow site – optimal based on Nash and Sutcliffe, 1970)	Yes – London WRZ
Wilby & Harris, 2006	1 – Thames to Kingston	2020s/ 2050s/2080s	Yes – weights applied to each (equal weights for emission scenarios and downscaling)	2 SRES (A2 and B2)	4 – CGCM2, CSIRO Mk2, ECHAM4 HADCM3	None	2 methods – Perturbed baseline with monthly change factors and SDSM	Not reported	3 HRUs  Areal representative climatic data	Daily flow sequences	2 – CatchMod and REGMOD	100 most skilful for CatchMod (based on Nash-Sutcliffe (1970)). Multiple sets for REGMOD (randomly sampled)	No
Wilby, 2005	1 – Thames to Kingston	2020s/ 2050s/2080s	No	2 SRES (A2 and B2)	1 –HADCM3	None	Statistical downscaling model (SDSM)	MORECS Penman-Monteith (Hough & Jones, 1997)	3 HRUs  Areal representative climatic data	Daily flow sequences	1 – CatchMod	Various combinations taken from ensemble of 10,000 randomly sampled sets	No
Wilby et al., 2006	1- River Kennet	2020s/2050s/ 2080s	No	2 SRES (A2 and B2)	3 (HADCM3, CGMC2, CSIRO MK2)	None	Statistical downscaling model (SDSM)	MORECS Penman-Monteith (Hough & Jones, 1997)	Spatiality of climatic data not reported  4 HRUs	Daily flow sequences	1 – CatchMod	1 – set (objective function(s) not explicitly reported	No
Wilby et al., 2011	1 – River Itchen	2020s	No but sensitivity testing used	1 SRES A1F1	Based on UKCP09 Weather Generator  2 ensemble members used  No bias-correction			MORECS Penman-Monteith (Hough & Jones, 1997)	2 - HRUs	Daily flow sequences	1 - CatchMod	1 – manual calibration for each HRU based on Nash and Sutcliffe, 1970	No but licensing impacts considered





### 2.3.1 *Water Resources industry focussed studies assessing the impacts of climate change*

As part of an UKWIR (UK Water Industry Research) sponsored study, Prudhomme et al (2005) considered the potential significance of a range of uncertainty sources to provide guidance to the UK water resources industry on planning for the impacts of climate change. Prudhomme et al. (2005) considered two emission scenarios, three GCMs and three downscaling techniques (Statistical [SDSM and delta change methods] and dynamical), two versions of the PDM model structure (one considered more ‘physically-based’ than the other) and multiple hydrological parameter sets (with both structures adopting a lumped form) to explore changes in river flows in thirteen British catchments under climate change. The 2020s time-horizon was the main focus of the study with some assessment also undertaken for the 2080s. Multiple performance measures (including volume-error [see Gupta et al., 1999] and Nash-Sutcliffe efficiency, 1970) over split calibration / validation periods were used to identify an ensemble of satisfactory models for each catchment. Two conceptual (i.e. not real) water resource systems were modelled to assess the impacts on system performance. Prudhomme et al. (2005) concluded that GCMs models showed the largest uncertainty in river flow change compared to the downscaling approach and emissions uncertainty but that hydrological uncertainty (structural and parametric) was significant with the size of the uncertainty range depending on the model used, and could vary significantly from one catchment to another.

To derive regional scale factors for use in estimating the potential impacts of climate change on river flows and groundwater, a follow-on UKWIR (2007) sponsored project developed a framework based around the projections from six climate models. This study used two hydrological models (PDM and CatchMod), considered alternative downscaling methods and the sampling of multiple hydrological parameter sets to translate climate changes at the global scale to changes in river flow and groundwater recharge (at the catchment scale) for 70 UK catchments. The framework developed a selection scheme to exclude or “downweight” ensemble members that failed to meet specific performance criteria (e.g. the degree of association of model predictions to observations for climate features of interest) using a Generalised Likelihood Uncertainty Estimation (GLUE) methodology (Beven & Binley, 1992). Whilst the methodology applied in this UKWIR study was relatively simple and highlighted non-linear responses of a system, the approach did have limitations as acknowledged by the study’s authors. The use of perturbation factors for exploring climate change did not take into account any changes in weather sequences and therefore future droughts had identical temporal patterns as the observed droughts. In addition, the study considered a limited number of parameter sets (relative to the number of parameters to be calibrated), applied uniform climatology to all

catchments (with the largest almost 10,000 km<sup>2</sup>), used a single performance measure (the Nash-Sutcliffe (1970) efficiency measure) and provided limited evidence as to the 'skill' of the resulting ensembles in prediction upon which all such design choices (model structure, parameter sampling, performance measure, spatial discretisation of climate and model) might be expected to have an influence. The output from this study provided a methodological approach, with associated regional scale factors, for the UK water resources industry to account for climate change in the water resources planning process based on the best available evidence at that time.

The Future Flows project (Prudhomme et al., 2012) produced an eleven-member ensemble, nationally consistent, projection of 148 years (1951-2098) of daily river flow and monthly groundwater level time series for 282 river catchments and 24 boreholes in Great Britain. These outputs were generated using the Future Flows climate projections (see Section 2.2.2.2) as the forcing climate data. The three model structures (CERF (see Griffiths et al., 2008), PDM and CLASSIC) used for the specific catchment modelling of river flows employed three different methods of calibration with the emphasis of calibration on different parts of the flow regime. For CERF the emphasis was on water resources as represented by the water balance and low flows, while for PDM and CLASSIC the emphasis was on the upper part of the flow regime and peak flows. The report states that the calibration method may affect model performance at different parts of the flow regime.

For four contrasting catchments, the Future Flows outputs were compared with simulations based on the 10,000 UKCP09 probabilistic climate change scenarios. This comparison suggested that the Future Flows outputs generally capture most of the range associated with simulations using the UKCP09-probabilistic ensemble but did not describe the upper tail of the distribution with the Future Flows projections lying in towards the lower half of the distribution from using the UKCP09 scenarios. The Future Flows project also highlighted the assumption in such studies that hydrological model parameters are fixed (a single parameter set was used for each catchment in this research project) and therefore assume, as with most similar studies, a stationary catchment (i.e. the climate may change but how the catchment responds to the inputs of precipitation and PET does not).

The hydrological modelling outputs from this project were incorporated as a source of evidence into the current water resources planning guidance (Environment Agency, 2012; 2017) for potential application by water companies in England and Wales. However, despite the ambition of the Future Flows project to provide scenarios appropriate for hydrological and hydrogeological application, there is limited evidence, at the time of this thesis, in the context of water resources

planning where this ambition has been examined outside of the Future Flows project itself. Huskova et al. (2016) did use the Future Flows projections in their research into water resources infrastructure investments in London but the Future Flows projections themselves were not examined in any detail. Thames Water (2014) also made use of the Future Flows projections in their Water Resources Management Plan but published only a limited commentary associated with this analysis to provide evidence as to their appropriateness for water resources applications. To address this knowledge gap, the Future Flows projections are described in more detail and examined in Chapter 5 to evaluate their potential suitability for water resources planning.

### *2.3.2 Research into the climate change impacts for water resources planning*

This section aims to provide an overview of recent academic research into the supply-side aspects of water resources planning under uncertainty. This section highlights the methodological features, key conclusions and identified knowledge gaps from this research that are considered particularly relevant to this thesis. These are presented under the following five topics:

- Approach to hydrological modelling uncertainty;
- Application of probabilistic climate change projections;
- Estimation of potential evapotranspiration;
- Quantifying impacts on water resources systems;
- “Top-Down” versus “Bottom-Up” approaches.

#### *2.3.2.1 Approach to hydrological modelling uncertainty*

The performance of any hydrological model is influenced by the choice of model structure and spatial complexity (i.e. the underlying equations, concepts and numerical schemes used along with how a catchment is discretised), the approach to evaluating model performance and the availability and accuracy of the input data (e.g. Beven 2001; Clark et al. 2008; Westerberg et al. 2011b; Krueger et al. 2010; Beven & Westerberg, 2011). This section aims to provide an overview of recent research that has explicitly considered hydrological modelling uncertainty in the context of water resources planning.

The Thames catchment, which is the case study used in this thesis, has been the subject of significant previous research in water resources planning under uncertainty (e.g. Wilby, 2005; Wilby & Harris, 2006; New et al., 2007a; Manning et al., 2009; Matrosov et al., 2013; Borgomeo et al., 2014; Walsh et al., 2015). The majority of this research has used the CatchMod hydrological model structure (e.g. Wilby, 2005; Wilby & Harris, 2006; New et al., 2007a; Manning et al., 2009;

Matrosov et al., 2013; Borgomeo et al., 2014; Walsh et al., 2015) to simulate river flows. With the exception of two studies (Wilby & Harris, 2006; Matrosov et al., 2013) the majority of this research did not compare the performance of CatchMod with alternative model structures.

Matrosov et al. (2013) made use of the PDM model structure, in addition to CatchMod, but did not provide a commentary as to their respective performance in simulating flows in the River Thames. Wilby & Harris (2006) used a multiple linear regression model (REGMOD), along with CatchMod, to provide an assessment of hydrological model structure uncertainty. Wilby & Harris (2006) reported systematic differences between CATCHMOD and REGMOD and suggested that the more conservative response from CATCHMOD may have been a reflection of the longer “memory” of wet winter conditions that was not captured by the multiple linear regression model.

Previous research using CatchMod to model the River Thames (e.g. Wilby, 2005; Wilby & Harris, 2006; New et al., 2007a; Manning et al., 2009; Matrosov et al., 2013; Borgomeo et al., 2014; Walsh et al., 2015), has not typically included an evaluation of its ability to reproduce the key features which could impact the performance of the public water supply system. One exception, Wilby & Harris (2006), did reflect on abstraction for public water supplies and reported that whilst CatchMod performed well at low to medium flows, it demonstrated a tendency to simulate too many days (compared to the observed record) above a threshold which would impact abstraction. Wilby & Harris (2006) also highlighted limitations with CatchMod’s ability to reproduce high flow periods.

Away from modelling the River Thames, CatchMod was used by Lopez et al. (2009) in their study of water resources for the Wimbleball area (south-west of England). Lopez et al. (2009) adopted an existing calibrated CatchMod parameterisation and highlighted that this single model realisation was not able to reproduce low flows particularly well compared to the observed record. Cloke et al. (2010) also applied the CatchMod model to investigate climate change impacts in the River Medway catchment. Cloke et al. (2010) reported that whilst CatchMod was considered suitable for investigating climate change impacts, low flows were consistently found to be under-predicted (particularly in spring and summer). Cloke et al. (2010) suggested that this under-prediction might be due to model structural limitations or possibly due to uncertainty regarding how the gauge record was naturalised and historical abstractions and discharges taken into account. CatchMod (along with PDM) was also used by Christensen et al. (2012) in their modelling of 70 UK catchments using the UKCP09 projections although this study did not examine its relative performance in detail for each of the studied catchments.

In other water resources research not specifically focussed on the Thames catchment, a number of alternative hydrological model structures have been adopted. These include the PDM (Prudhomme & Davies, 2009; Christiersen et al., 2012; Matrosov et al., 2013; Charlton & Arnell, 2014) and ARNO (Fowler et al., 2008) model structures. Similarly to the above research using CatchMod, a critical examination of these model structures' ability to reproduce the features which directly influence the performance of public water supply systems was typically not included as part of the published research.

A repeated commentary across the majority of the research discussed above is that multiple model structures should be considered as part of future research to provide a more rigorous consideration of hydrological modelling uncertainty (e.g. Manning et al., 2009; Borgomeo et al., 2014), mitigate biases and limitations related to a specific model structure (e.g. UKWIR, 2007) and ultimately lead to improved projections of river flow under future climate (e.g. Cloke et al., 2010).

Another key aspect of hydrological modelling uncertainty considered in previous research is hydrological model parameter uncertainty. To consider parameter uncertainty, a number of recent water resources related studies have adopted a GLUE (Beven & Binley, 1992 – see Section 2.3.2) based approach (e.g. Cloke et al., 2010; Christiersen et al., 2012; Matrosov et al., 2013; Borgomeo et al., 2014).

Cloke et al. (2010) reported that CatchMod, applied as part of a GLUE based approach, was useful for considering hydrological model parameter uncertainty in projecting the impacts of climate change on river flows. Cloke et al. (2010) used the Nash-Sutcliffe efficiency metric on flows (and the natural logarithm of flows) to evaluate model performance. The resulting likelihood-weighted flow distributions were reported as capturing the variation in observed low and medium flows sufficiently well. Through examination of the parameter sets considered behavioural under the GLUE approach, Cloke et al. (2010) identified specific individual parameters that significantly influenced either low or high flow responses and suggested that the significance of parameter uncertainty increased at lower flows.

Christiersen et al. (2012) used the Nash-Sutcliffe efficiency criterion (Nash & Sutcliffe, 1970) on the natural logarithm of flows, as part of a GLUE analysis, to select parameter sets from an initial sample size of 5,000 for each of two model structures. Parameter sets were retained, for each of 70 UK catchments, where a specified threshold over a baseline period of 1961 to 1990 was achieved. Christiersen et al. (2012) reported that the uncertainty associated with climate change projections dominated systematically over the uncertainty in hydrological modelling response (when considering the 2020s time-horizon using the UKCP09 probabilistic projections). However,

it should be noted that this study only considered changes in monthly mean flows rather than specifically looking at prolonged periods of low flow.

In their water resources study of London, Matrosov et al. (2013) reported that hydrological model parameter sets were chosen based on how closely simulated flows matched the historical record. However, the metric used to select and weight parameter sets as part of their GLUE-based methodology was not explicitly stated. Further, despite carrying forward multiple parameter sets, Matrosov et al. (2013) did not discuss the relative influence of hydrological modelling uncertainty as part of their study on alternative decision-making methods.

Borgomeo et al. (2014) considered hydrological model parameter uncertainty through identifying 53 CatchMod model realisations from an original Monte Carlo sample of 10,000. These 53 parameter sets were chosen based on a volume error performance criteria assessed at a single location over a single time-window and was used in prediction adopting a GLUE-based approach. Borgomeo et al. (2014) reported that hydrological model parameter uncertainty had a significant impact on the simulation results, particularly when considering the impacts on the reliability of public water supplies in London and therefore should not be ignored.

Other studies, whilst not adopting a GLUE based approach, have also used multiple parameter sets to reflect hydrological parameter uncertainty. Wilby (2005) derived 10,000 alternative CatchMod hydrological parameter sets, considering different subsets of training data, to explore the uncertainty in the hydrological modelling process. Hydrological model performance was evaluated using two performance metrics (Nash-Sutcliffe efficiency criterion and Relative Absolute Mean Error, RAME) to select 100 parameter sets. These sets were subsequently used to consider the sensitivity of impacts to hydrological model parameterisation. As part of the research, Wilby (2005) reported that hydrological parameter uncertainty had a significant influence on model predictions and demonstrated the relative performance of different parameter sets during contrasting periods of the historical flow record.

Manning et al. (2009) built on the research of Wilby (2005) and identified eleven new parameterisations of the same CatchMod model using similar performance metrics. These eleven parameterisations were considered Pareto-optimal, meaning that for each parameterisation, no improvement in agreement with one criterion could be made without compromising the other. These eleven parameterisations were used to explore the sensitivity of model predictions to parameter uncertainty and Manning et al. (2009) reported that the influence of hydrological parameter uncertainty was significant, particular for estimating the impacts of climate change on low flows.

Wilby & Harris (2006) also used the Nash-Sutcliffe efficiency criterion, as a single performance metric by which to select the 100 best sets of hydrological model parameters. These selected parameter sets were subsequently weighted based on their relative performance and used to reflect hydrological modelling sensitivity in projecting climate change impacts. As part of this research, Wilby & Harris (2006) commented that replacing the Nash-Sutcliffe efficiency criterion with a measure of absolute mean error had a negligible effect on river flow projections. The same parameter sets of Wilby & Harris (2006) were used by New et al. (2007a) in their water resources study for the Thames. New et al. (2007a) reported a significant influence due to parameter uncertainty on projected impacts of climate change on river flows, particular at low flows.

The approach to parameter uncertainty adopted by Prudhomme & Davies (2009) took an ensemble of 'near optimal' parameter sets, with performance measured using the Nash-Sutcliffe efficiency measure (1970). Prudhomme & Davies (2009) reported that the relative uncertainty due to hydrological parameterisation compared to GCM and emission scenario was low but its significance varied across different catchments and should be considered as part of future water resources studies.

Charlton & Arnell (2014) calibrated a form of the PDM model for six UK catchments using Nash-Sutcliffe efficiency criterion, volume-error and visual inspection of flow duration curves over split calibration-validation periods. In this research, parameter uncertainty was considered by allowing each parameter to vary by up to plus or minus 10%, with each perturbed parameter set sampled across all five parameter spaces independently. Volume error and Nash-Sutcliffe indices were calculated for the calibration period for each set of perturbed parameters and 100 additional parameter sets considered to produce "good" fits (where volume error was within 5% of that achieved by the calibrated parameter set) were also taken forward. The influence of hydrological parameter uncertainty varied by catchment but was considered less significant than the uncertainty due to GCM when considering projected changes in climate forcing – as indexed by change in global average temperature, from 0.5 to 6 °Celsius.

Other water resource focussed studies of interest have only used a single, considered optimal parameter set. Fowler et al. (2007b) used the Nash-Sutcliffe efficiency criterion to calibrate their hydrological model for generating inflows for their water resources system modelling in the Integrated Resource Zone in the north-west of England. Similarly, for their study of both high and low flows in the River Eden, Fowler et al. (2008) used both the Nash-Sutcliffe efficiency criterion and the overall water balance in the hydrological model to identity a single, optimal parameter set. Lopez et al. (2009) made use of an existing calibrated CatchMod model as part of their water



resources modelling study in the south-west of England. Walsh et al. (2015) in their study of the London, used a single hydrological model realisation, taken from the research by Manning et al. (2009), for generating inflows into their water resources model.

Finally, Clark et al. (2016) emphasised the importance of assessing the uncertainty from a number of sources in developing projections of water availability and reviewed the current capabilities and limitations for characterising and understanding uncertainty in the hydrological impacts of climate change in this context. Clark et al. (2016) advocated improved characterisation of uncertainty in hydrological modelling, using frameworks designed to accommodate multiple spatial configurations, multiple process parameterisations, and multiple model parameter values along with reducing hydrological model uncertainty through advances in process representation.

#### 2.3.2.2 Application of probabilistic climate change projections

Wilby & Harris (2006) presented a probabilistic framework for combining information from four GCMs, two emission scenarios, two statistical downscaling approaches, two hydrological model structures, alternative hydrological parameter sets for three time-horizons (2020s, 2050s, 2080s). The climate change projections were weighted according to an index of reliability for downscaled effective rainfall, considered appropriate for water resources planning in the Thames catchment. Wilby & Harris (2006) reported that the framework demonstrated how components of uncertainty could be objectively weighted, leading to conditional probabilities for climate change impact assessments. Wilby & Harris (2006) also highlighted the need for further work to develop practical guidance for planners and engineers who have the difficult task of translating probabilities into adaptation responses.

New et al. (2007a) also explored the implications of probabilistic end-to-end risk based frameworks for climate impacts. Monthly change factors were derived from a large (2,700-member) ensemble of climate change projections using a single GCM model structure. New et al. (2007a; 2007b) suggested that a probabilistic approach potentially provides more informative results than scenario-based approaches for the decision-maker. However, to move beyond their illustrative example New et al. (2007a) highlighted that more elements would need to be considered. These include consideration of downscaling and hydrological model structural uncertainties, a more sophisticated approach to assessing and weighting the skill of individual model combinations and the use of a water resource systems model to enable the assessment of the interplay of demand and supply under different socio-economic and water infrastructure scenarios.

Fowler et al. (2008) explored the development of probabilistic estimates of climate change impacts on river flows in the River Eden, Cumbria. Thirteen RCMs, driven by boundary conditions from two GCMs, were used to generate Probability Density Functions (PDFs) of change in temperature and precipitation for the 2020s, 2050s and 2080s. A weather generator was used to downscale these PDFs to the catchment scale, using change factors calculated as the difference between the future and control time-periods within the RCM simulations. The results from the different RCMs were weighted using a Bayesian method with the weight applied to a selected RCM a function of its performance in reproducing current climate (1961 to 1990) and its agreement with the ensemble consensus for future projections (which assumed independence between ensemble members). The results from the modelling were a set of PDFs of river flow statistics that provided a probabilistic assessment of climate change impacts. Fowler et al. (2008) concluded that potential improvements to the approach may include an assessment of the future uncertainty introduced by the emission scenario, and the structure and parameterisation of the hydrological model.

Lopez et al. (2009) evaluated the value of perturbed physics ensembles (CPDN, 2016) of climate models for understanding and planning changes to public water supply, over decadal time-periods to the 2070s. Lopez et al. (2009) concluded that the additional information contained in a climate model ensemble provides a better understanding of the possible ranges of future conditions, compared to the use of a single climate model scenario. However, results were not weighted by Lopez et al. (2009) because of concerns in attempting to quantify the different scales of uncertainties.

Christierson et al. (2012) built on the work undertaken by UKWIR (2007) and used the modelling framework developed in that study to explore the use of the probabilistic UKCP09 climate projections. This study only considered the 2020s time-horizon and a single emission scenario for assessing the impacts of climate change on river flows in the context of water resources planning. Christierson et al. (2012) demonstrated the use of Latin Hypercube Sampling (McKay et al., 1979) to identify a representative sub-sample of climate projections (from the 10,000 member UKCP09 ensembles) to reduce computational demands and encourage subsequent adoption of the UKCP09 projections in the wider water resources industry. For the Thames catchment, Christierson et al. (2012) reported that a LHS sample size of 20 captured most of the uncertainty in terms of flow impacts in the River Thames.

Charlton & Arnell (2014) demonstrated the use of the full 10,000 member UKCP09 projections (for three emission scenarios and three time-horizons) on six UK catchments. In addition, Charlton

& Arnell (2014) compared the use of the full ensemble along with different randomly sampled subsets ranging from 10 to 1,000 on one of the UK case studies. This research focussed on both high and low flows (with low flows represented by the Q95 flow – the flow exceeded 95% of the time). The authors reported a large range in hydrological changes across the six study catchments under the UKCP09 climate projections, with the differences between catchments primarily due to differences in catchment geology and the baseline water balance. In terms of sampling of the UKCP09 projections, Charlton & Arnell (2014) reported that with randomly sampled subsets of 10 scenarios both the range and the distribution of changes varied considerably. With 20 randomly sampled scenarios, Charlton & Arnell (2014) reported that there was more consistency in the range of changes, but the shape of the distribution could still be very different between different subsets. With at least 100 randomly sampled scenarios, Charlton & Arnell (2014) reported much greater consistency in both the range and shape of the distributions.

As part of their risk-based approach to water resources planning, Borgomeo et al. (2014) considered 10,000 realisations of transient sequences between 2001 and 2060, along with two population change forecasts to reflect projected changes in water demand. The transient climate realisations were generated using a stochastic weather generator, with inputs informed from examination of the trajectory of changes in climate variables suggested by the probabilistic UKCP09 projections.

Walsh et al. (2015) used a rainfall generator in combination with 100 randomly sampled scenarios from the UKCP09 probabilistic projections. In this research, a spatially coherent rainfall generator was used, intended to help capture non-linear impacts of climate change on water resources. 100 realisations were considered for each time-horizon and Walsh et al. (2015) provided a supplementary note discussing the validation of the rainfall generator based on monthly rainfall statistics and comparison with the historical record. However, prolonged drought periods of particular interest to water resources planners were not considered within this supplementary note.

As a contrast to the “top-down” studies (see section 2.3.2.5) described above, Turner et al. (2014) evaluated the risks posed by climate change using a “bottom-up” decision-scaling approach (Brown et al., 2012). This research did not make use of hydrological models to provide inflows and, instead, used synthetically generated river flow sequences for multiple sites in Melbourne, Australia, to reflect the uncertainty in future flows. Consequently, no evaluation of the uncertainties associated with hydrological model (structure or parameter) or estimation of weather metrics was undertaken as part of this research. The uncertainty with regards to climate

change was based on 138 climate model simulations using 23 climate models run under six emissions scenarios developed as part of the CMIP3 study.

### 2.3.2.3 Estimating potential evapotranspiration

In recent research into water resources planning, only a single method of estimating PET has typically been used. These methods include the Penman (1948) equation (e.g. New et al., 2007a), the Penman-Monteith (e.g. Allen, 1998) method (e.g. Wilby & Harris; 2006, Prudhomme & Davies, 2009; Manning et al., 2009; Lopez et al., 2009; Cloke et al., 2010), the Oudin method (e.g. Christerson et al., 2012; Matrosov et al., 2013) and the Blaney-Criddle (1950) formula (e.g. Fowler et al., 2007b; Fowler et al., 2008).

In their water resources study on the Thames catchment, Manning et al. (2009) calculated PET using the MORECS method as part of using a weather generator (Hough & Jones, 1997) to generate rainfall and PET sequences under a changed climate. Manning et al. (2009) highlighted that the weather generator was not able to reproduce PET well for the hot, dry summer of 1976. Manning et al. (2009) suggested that the possible underestimation of PET in very hot summers is partially mitigated in the subsequent hydrological modelling because actual evaporation is limited by moisture supply rather than determined by PET. Manning et al. (2009) demonstrated the sensitivity of river flows to PET estimates and emphasised that correct calculation of PET is essential in calculating future water resource availability.

Christerson et al. (2012) included a discussion on the uncertainty due to estimating PET and highlighted that the issue of PET formulation for climate change impact assessments is currently far from settled. In their study, Christerson et al. (2012) used only the Oudin method to reduce computational resource requirements and highlighted that choosing another formulation of PET would have inevitably changed the sets of behavioural models for each catchment, making uncertainty assessments and comparisons rather problematic.

For their research using the UKCP09 probabilistic projections, Charlton & Arnell (2014) estimated PET using the Penman-Monteith method and assumed that wind speed was unchanged under a future climate (noting that under UKCP09, changes in wind speed have since been provided as a separate product but cannot be used in conjunction with other UKCP09 variables - see UKCP09 (2017)). Charlton & Arnell (2014) suggested that projected changes in PET due to climate change are sensitive to the formulation used, and this may be most significant at low flows. Further, Charlton & Arnell state that appropriate PET estimation could be particularly important in regions where precipitation and PET are in close balance. Finally, Charlton & Arnell (2014), based on Bell

et al. (2011), suggest that the inclusion of potential changes in crop parameters under higher carbon dioxide concentrations could also make a difference to projected changes in PET.

In two recent studies exploring water resources planning under uncertainty (Borgomeo et al., 2014; Walsh et al., 2015) the PET method was not reported. Walsh et al. (2015) reported that a single record of PET at the centroid of the Thames basin was adopted and, although not explicitly reported, it is assumed this was also applied to the Lee catchment as part of the water resources system modelling.

#### 2.3.2.4 Quantifying impacts on water resources systems

Fowler et al. (2007) considered the impact of climate change on the operation of a complex conjunctive-use water supply system in north-western England. In this study, impacts were quantified based on comparison of reservoir storage, estimated yield (the maximum demand that could be placed without causing any deficits) and the number of days during which drought restrictions were triggered. Drought restrictions in this system were triggered when the combined reservoir storage dropped below a specific threshold. Similarly, Lopez et al. (2009), in a case study undertaken on the Wimbleball water resources zone in the south-west of England, quantified impacts in terms of changes to reservoir storage and frequency of failure to fully satisfy demand.

New et al. (2007a) assessed potential impacts on water resources in London using the frequency of target flow thresholds for the River Thames being satisfied. New et al. (2007a) did not include water resource system modelling and highlighted that demand restrictions in London are triggered by both flow and reservoir storage thresholds (see Section 6.3.4.4 for a detailed description of these thresholds) with the latter not considered as part of their study. Manning et al. (2009) did include water resources system modelling and quantified the implications on water resource system performance in terms of the changing potential in abstraction from the River Thames based on flows in the River Thames only.

Whilst not directly modelling water resource system impacts, Wilby et al. (2011) employed the CatchMod model structure and two climate change ensemble members, using the UKCP09 weather generator, as part of a modelling framework for evaluating the sensitivity of low river flows to different configurations of abstraction licensing under both historical climate variability and expected climate change. Using the River Itchen in southern England as a case study, Wilby et al. (2011) showed that the abstraction volume was more sensitive to uncertainty in the regional climate change projection than to the environmental flow target. Wilby et al. (2011) also found that natural climate variability over multi-decadal time scales was shown to be a significant factor affecting the frequency of low flows and abstractable volumes. Even 20<sup>th</sup> century variability in the

precipitation regime would have caused the number of days with low flows to vary by an order of magnitude between different 30 year periods.

Matrosov et al. (2013) assessed impacts on public water supplies using a service reliability criterion based on the frequency of water use restrictions imposed in the Thames basin. These restrictions corresponded to sprinkler and non-essential use bans, with average service reliability calculated by assessing the number of weeks each restriction level was imposed. A storage susceptibility metric was also defined as the lowest storage level reached by the combined reservoir storage for London with a specified threshold below which failure was triggered due to pressure-related distribution problems in the network.

Similarly to Matrosov et al. (2013), Borgomeo et al. (2014) and Walsh et al. (2015) also considered the frequency of demand restrictions as the metric by which to quantify the impacts on the performance of the public water supply system. This metric was considered to be of direct relevance to water resource planners because these Levels of Service (see Section 6.3.4.4) are what the performance of the water company is compared against.

Turner et al. (2014), in their study considering climate change impacts in Melbourne, Australia, used two service-based criteria: a minimum reliability criterion based on a water use intervention threshold defined by the Melbourne water companies and a vulnerability criterion based on a maximum allowable drawdown in total system storage. In effect both trigger points were based on a total system storage (TSS) with the reliability criterion set at a higher storage level compared to the vulnerability criterion.

Finally, when assessing the performance of a water resource system, the role played by the physical processes that occur during the end of a drought also warrant consideration. Parry et al. (2016a) presented a systematic assessment of drought termination in the UK and suggested that prolonged drought development phases tend to be followed by shorter and more abrupt drought termination and that drought terminations show considerable spatio-temporal variability. In a related study, Parry et al. (2016b) reported that the extent to which climate and hydrological models are able to simulate observed drought termination events remains an open question and that improved understanding of the physical processes during and at the termination of droughts may support refinement of hydrological model structures and lead to improved water resources planning and forecasting.

#### 2.3.2.5 “Top-Down” versus “Bottom-Up” approaches

The approaches used for evaluating and dealing with climate change may be characterised as being either “top-down” or “bottom-up” (Dessai & Hume, 2004; Carter & Mäkinen, 2011). The current WRMP climate change methodologies are a “top-down” approach and, although the source of climate change projections has changed as climate science and available projections have evolved, the principles of the planning methods have remained largely unchanged from the earliest UK climate change and hydrological impact studies (e.g. Arnell and Reynard, 1996).

“Top-down” approaches are considered to be “prediction-orientated” (Dessai & Hume, 2004) that, in the context of water resources, are based on a set of climate projections being simulated using hydrological, hydrogeological and water resource system models to determine the system’s performance with respect to each projection (Culley et al., 2016). The system’s performance is then classified as “acceptable” or “unacceptable” for each projection, and the potential benefits of alternative adaptation strategies can be explored (Prudhomme et al., 2010).

The principal challenge of using “top-down” approaches is that the assessments may only be valid for the climate models used in the analyses. This means that any assessment is only valid until new climate models are developed and can only cover the range of climatic futures covered by the available climate model(s). Due to the discrete nature of the projections used, such “top-down” approaches are generally not suitable for identifying thresholds of performance with respect to changes in climate exposure, as it may be difficult to identify the exact amount of climate change at which system performance changes from acceptable to unacceptable (e.g. Culley et al., 2016).

In terms of decision-making and adaptation planning, Clark et al. (2016) suggested that for many users, including water resources planners, climate change evidence is most compatible with the

decision-making processes when it is distilled into a set of discrete quantitative hydrological storylines of climate change impacts, each representing key features from the full range of possible climate scenarios. This echoes the commentary from Whetton et al. (2012) which, having noted that some providers of climate risk information have reappraised the ways in which climate model output is delivered to “users”, proposed the use of Representative Climate Futures (RCFs).

These RCFs are designed as a way of balancing the considerations of uncertainty and simplicity of scenarios for adaptation planners and are developed in the form of a small set of scenarios with descriptors (such as “slightly warmer with little rainfall change” or “hotter and drier”) with relative likelihoods given by climate model ensembles. The RCFs are then subsequently used as a framework in which to classify more detailed information, such as available climate model and downscaled data sets. Whetton et al. (2012) argued that since the RCF descriptions need not change as new climate change evidence emerges, they can provide a stable framework for assimilating risk assessments undertaken at different times with different sets of climate models. Whetton et al. (2012) did acknowledge that the approach does require various significant challenges to be addressed; in particular, robustly classifying future regional climates into a small set and estimating likelihoods.

Building on these concepts, and in contrast to the predominantly “top-down” approach to climate change adopted by water resource planners in the UK, a number of recent studies (e.g. Culley et al., 2016; Yates et al., 2015; Turner et al., 2014; Amarasinghe et al., 2016) have promoted “bottom-up” approaches.

An example of a “bottom-up” approach in the climate impact assessment arena is “decision scaling” (Brown et al., 2012), where a scenario neutral climate space is divided into regions for which different discrete decisions would be preferable, thus allowing the articulation of preferred adaptation options in response to specific changes in climate (Culley et al., 2016). Culley et al. (2016) consider this is a convenient approach as it:

- provides an understanding of system vulnerability,
- identifies decision thresholds that can be compared easily with climate predictions, and
- demonstrates whether a particular decision can achieve acceptable performance under given climate conditions.

Culley et al., (2016) built on the decision scaling theory by proposing a “bottom-up” approach to designing optimal feedback control policies for a water system in Lake Como, Italy, exposed to a



changing climate. This approach not only described optimal operational policies for a range of potential climatic changes but also enabled an assessment of a system's upper limit of its operational adaptive capacity, beyond which upgrades to infrastructure become unavoidable.

Yates et al. (2015) also suggested that the conventional "top-down" approach to providing advice for adaptation planning is poorly suited to the task. Yates et al. (2015) argued that planners are then left with an intractable range of possibilities, and may habitually resort to "low regret" decisions. Yates et al. (2015) promoted a systematic risk-management approach to adaptation planning that focuses on identifying and reducing vulnerabilities to a plausible range of climate scenarios, while maintaining the flexibility to respond to evolving conditions. In their study, Yates et al. (2015) adopted an approach described as a "multi-step decision support process" to efficiently explore climate change adaptation planning in Colorado. This approach was based on stochastic generation of plausible futures to sensitivity test the system to reveal non-linear or threshold behaviours to the climate-forcing and then interpreting the climate change evidence in this context.

In another study, Turner et al. (2014) evaluated the risks posed by climate change using a "bottom-up" decision-scaling approach (Brown et al., 2012) that identified thresholds at which interventions in a water resource system would be necessary. Climate change evidence, taken from the third Coupled Model Inter-comparison Project archive (CMIP3, Meehl et al., 2007) was then evaluated in this context to provide an assessment of system risk. Amarasinghe et al. (2016) also undertook a "bottom-up" assessment of a supply system in Australia. Their focus was on evaluating the resilience of the Southeast Queensland (SEQ) Water Grid system; a large system comprising numerous treatment plants and 12 reservoirs. Amarasinghe et al. (2016) argued that there are operational benefits to resilience assessment over more traditional impact-style assessments; if the conditions that are needed in order for the system to function are understood, then it is less necessary to be able to understand with certainty what the future conditions will be.

"Bottom-up" climate change approaches have not been widely used in industry in the UK with the exception of assessing the climate change sensitivity of a range of UK catchments to future flooding (Prudhomme et al., 2010). This "scenario-neutral" approach systematically tested a number of hydrological catchments to a range of climate parameters to develop a "response surface" of impacts and then subsequently overlaid climate change scenarios to identify projected risks. This approach was subsequently adapted for considering drought risk and explored for UK water resources as part of the Environment Agency project "Performance of water supply systems during mild to extreme drought" (Anderton et al., 2015). This developed a "bottom-up" resilience

focussed method for UK water supply systems with a focus on understanding system sensitivity to drought.

### 2.3.3 Chapter summary

The research questions posed in this thesis are set-out in Section 1.2. This literature review has provided an overview and background to the current state of climate change impact studies on water resources with particular reference to the uncertainties associated with hydrological modelling, estimating PET and climate change. The following research findings and knowledge gaps have been identified, with their relevance to the research questions indicated in parentheses:

#### ***Hydrological modelling uncertainty for water resources planning (Research Questions 1 and 4)***

- In recent research specifically looking at water resources planning under uncertainty (e.g. Borgomeo et al., 2014; Turner et al., 2014; Walsh et al. 2015), the uncertainty associated with the hydrological modelling process (including model structures, parameterisation and the method used to estimate PET) has received only modest scrutiny with limited evidence provided as to how well the underlying hydrological models reproduce the flow characteristics of most relevance to the water resources planner.
- Whilst it is recommended that multiple hydrological model structures are used (e.g. Wilby & Harris, 2006; New et al., 2007a; Cloke et al., 2010; Christerson et al., 2012; Borgomeo et al., 2014) the majority of previous water resources research (e.g. New et al., 2007a; Fowler et al., 2007b; Lopez et al., 2009; Matrosov et al., 2013; Borgomeo et al., 2014; Walsh et al., 2015) have only considered the use of a single hydrological model structure. In terms of the conceptual model structures used in previous studies, these have included CatchMod (e.g. Wilby & Harris, 2006; Lopez et al., 2009; Manning et al., 2009; Cloke et al., 2010; Matrosov et al., 2013; Borgomeo et al., 2014; Walsh et al., 2015), PDM (e.g. Prudhomme, 2005; UKWIR, 2007; Prudhomme et al., 2012; Matrosov et al., 2013; Charlton & Arnell, 2014) and ARNO (e.g. Fowler et al., 2007b; Fowler et al., 2008).
- Within the UK water resources planning industry water companies typically use a single conceptual hydrological model structure for modelling a catchment. The model structures currently used by water companies (e.g. Thames Water, Southern Water, United Utilities, Anglian Water, Severn Trent Water and Welsh Water) include CatchMod, HYSIM and PDM. However, it is not necessarily clear to water resources planners which models should be selected for a specific study with previous studies seldom including detailed evidence demonstrating the strengths and weaknesses of these model structures to

reproduce the decision-relevant flow characteristics of particular interest to the water resources planner.

- It is recognised that hydrological model parameter uncertainty can be significant for water resources studies (e.g. Wilby & Harris, 2006; Manning et al., 2009; Cloke et al., 2010; Borgomeo et al., 2014) with some studies suggesting this significance potentially increases at the extremes: high flows (e.g. Wilby, 2005; Westerberg et al., 2016) and/or low flows (e.g. Cloke et al., 2010; Tian et al. (2014); Westerberg et al., 2016). Borgomeo et al. (2014) further highlighted the significant influence that hydrological model parameter uncertainty may have on water resource system impacts by the 2050s.
- Whilst one industry study, UKWIR (2007), has explored the issue of parameter uncertainty, this aspect is not explicitly considered by water companies as part of their planning process (e.g. Thames Water, 2014)
- Euser et al. (2013) suggested that model evaluation based on single objective optimisation is insufficient to appropriately identify dominant processes and that the use of a multi-objective optimisation may offer a better approach for understanding and visualising the strengths and weaknesses of models. Broderick et al. (2016) also recommended adopting multiple performance criteria that are pertinent to the study objectives when assessing the transferability of model parameters between contrasting climates.
- The method used to estimate PET has, to date, received limited scrutiny (Prudhomme & Williamson, 2013) but recent research suggests its influence could be significant (Haxton & Young, 2012; Christerson et al., 2012; Seiller & Anctil, 2016), particularly in the context of water resources planning in permeable lowland catchments in the south of England (Charlton & Arnell, 2014).
- Clark et al. (2016) advocated improved characterisation of uncertainty in hydrologic modelling, using frameworks designed to accommodate multiple spatial configurations, multiple process parameterisations, and multiple model parameter values.

#### ***Climate change uncertainty for water resources planning in the UK (Research Questions 2 and 5)***

- it is anticipated that climate change is likely to have a significant impact on water resources across the United Kingdom (e.g. Wilby & Harris, 2006; New et al., 2007a; Fowler et al., 2008; Cloke et al., 2010; Christerson et al., 2012; Borgomeo et al., 2014) and any significant reduction in water availability, particularly in areas of the country considered to be already “water stressed” (Environment Agency, 2008), could have major implications for our quality of life and the environment.

- Current UK industry guidance (Environment Agency, 2012; 2017) recommends the use of two sources of climate projections, the UKCP09 (Murphy et al., 2009a; 2009b) ‘probabilistic’ projections and the eleven-member Future Flows (Prudhomme et al., 2012) transient climate ensemble, noting that both products are strongly related given their underlying climate models. These two contrasting climate change “products” offer alternative approaches to considering climate change but limited evidence has been published comparing their suitability for water resources planning or demonstrating how they might be practicably used as part of a wider study considering multiple sources of uncertainty.
  - The UKCP09 projections have been used in several industry and academic studies for exploring climate change impacts on future river flows (e.g. Cloke et al., 2010; Christiernson et al., 2012; Matrosov et al., 2013; Charlton & Arnell, 2014; Borgomeo et al., 2014; Walsh et al., 2015). However, it is not always feasible to apply all scenarios from such large ensembles as part of a wider uncertainty study. Practicable approaches to reducing the computational burden, through selecting a representative sub-sample (e.g. Christiernson et al., 2012; Charlton & Arnell, 2014; Whetton et al., 2012) or adopting a “bottom-up” approach (e.g. Yates et al., 2015; Culley et al., 2016) may be more appropriate.
  - With regards to the Future Flows projections, whilst these do potentially allow the testing of a system under weather sequences with different temporal sequencing to the historical record, there is currently limited published evidence as to how the extreme events of interest to water resources planners are reproduced. A further limitation is that the Future Flows projections comprise only eleven members and therefore can only reflect a subset of the known uncertainties (Prudhomme et al., 2012).
- Another source of uncertainty that has received limited attention to date (Prudhomme & Williamson, 2013) in the context of water resources planning under a changed climate, is the method used to estimate PET. Manning et al. (2009) and Haxton & Young (2012) highlighted that the method used to estimate PET for future climate can be significant and Charlton & Arnell (2014) emphasised that the differences due to the choice of method may be particularly significant at low flows which is of relevance to the water resources planning process.

### ***Water resources planning under multiple sources of uncertainty (Research Questions 3 and 6)***

- Previous research which has sought to quantify the relative significance of different sources of uncertainty has concluded that uncertainty due to modelling climate change scenarios is large relative to the uncertainty in the hydrological modelling process but that that latter is significant and should be included in impact assessments (e.g. Wilby, 2005; Wilby & Harris, 2006; New et al., 2007a; Manning et al., 2009; Prudhomme & Davies, 2009; Charlton & Arnell, 2014; Borgomeo et al., 2014).
- Only some of the studies which have considered hydrological modelling using climate change for water resources planning (e.g. New et al., 2007a; Lopez et al., 2009; Matrosov et al., 2013; Borgomeo et al., 2014; Walsh et al., 2015) extended their analysis to explicitly quantify the impacts on the performance of water resource systems and understand the consequences on the reliability of future public water supplies.

The following chapter introduces the case study catchments, along with a description of the hydrological modelling framework and the historical data used in this thesis.

## Chapter 3 Methodology

### 3.1 Introduction

The focus of the research presented in the remainder of this thesis is to understand and contrast the role of some of the uncertainties in the water resources planning process. This chapter discusses elements of the research which are common to all chapters, with additional detail subsequently provided in each chapter where necessary.

This first part of this chapter describes the case study catchments. This is followed by a description of the hydrological modelling framework, Kestrel – IHM (see Section 3.3), and the integrated hydrological model structures, CatchMod and PDM-AQ, used throughout this thesis. Finally, this chapter concludes with a discussion of the input data used including the two methods for estimating potential evapotranspiration considered in this research.

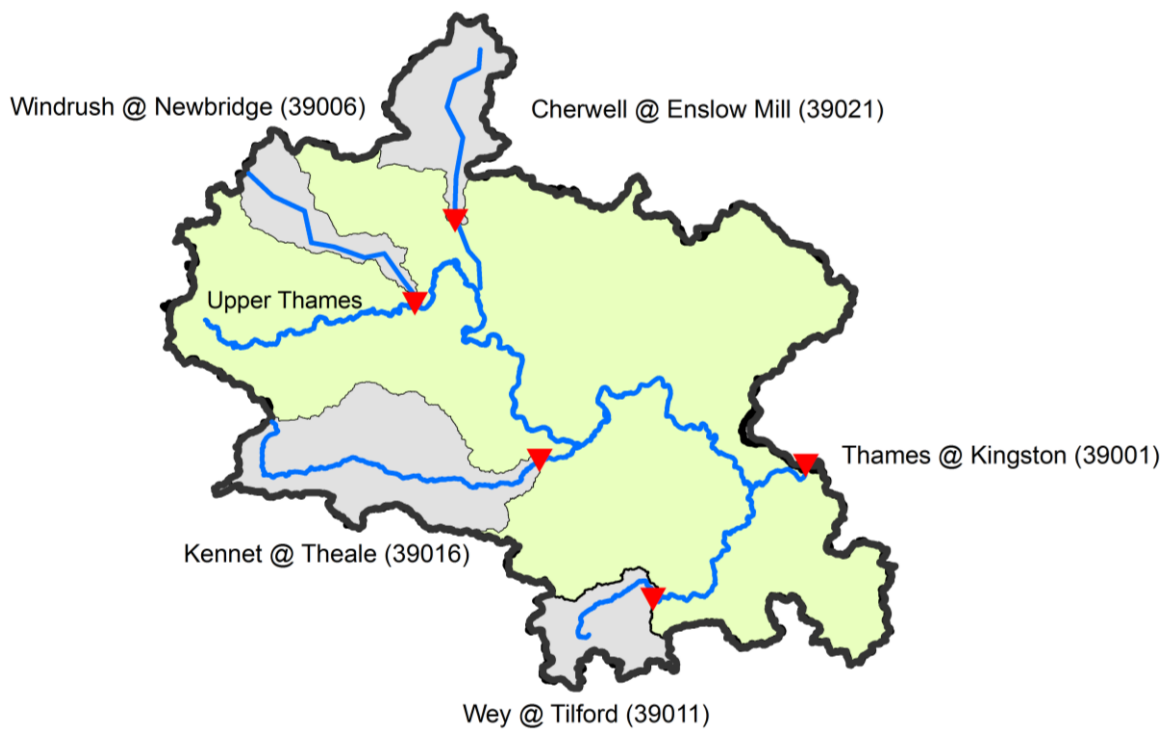
### 3.2 Description of the Thames catchment

The source of the River Thames is in the Cotswolds in Gloucestershire, with an area of approximately 9,950 km<sup>2</sup> draining to the river at Kingston and average annual rainfall varying from approximately 850mm to 600mm across the catchment (NRFA, 2016). As described by Bloomfield et al. (2009), the Thames basin is underlain by a thick sequence of Mesozoic to Recent rocks that can be divided into three broad structural zones based on geological structure: the Midlands Shelf to the north-west; the London Basin in the central area; and the Wealden Anticline to the south-east, each with their own characteristic lithostratigraphy. The major aquifer in the basin is the Upper Cretaceous Chalk, which is part of the Cretaceous and Palaeogene rocks within the Midlands Shelf. Palaeogene to Recent surficial deposits can be found throughout the Thames basin across all three of the structural zones.

The Thames River basin is of regional importance, supplying a number of water resource systems both through abstractions from the river system and also via groundwater abstractions from the underlying aquifers. One of these systems is Thames Water's London Water Resource Zone (WRZ) which covers the majority of the Greater London area, supplying a population of over 6.5 million and is particularly dependent upon abstractions from the River Thames which are then stored in large raw water reservoirs.

The modelling presented in this thesis considers the five locations within the Thames river basin shown in Figure 3-1. These have been chosen based on the availability of historical flow records, their geographical locations and contrasting hydrogeological characteristics (refer to Table 3-1)

along with their relevance to water resources planning which is described later in this thesis (see sections 4.4.2 and 6.3.1).



**Figure 3-1: Thames catchment showing the locations considered in this thesis.**

### 3.3 Development of a hydrological modelling framework for water resources planning

As highlighted in Chapter 2, Clark et al. (2016) advocated improved characterisation of uncertainty in hydrological modelling for water availability studies, using frameworks designed to accommodate multiple spatial configurations, multiple process parameterisations, and multiple model parameter values. For this research, an ensemble hydrological modelling framework (Kestrel – IHM) has been developed to allow multiple catchments to be modelled using two different model structures and allow for different levels of spatial complexity in the soil moisture accounting process to be considered. In addition, these models can be ‘driven’ by ensembles of climate inputs in a computationally efficient manner, automated using a scripting process.

The framework has been developed using Microsoft’s .NET software development framework and also makes use of the open source DotSpatial Geographic Information System ([www.dotspatial.org](http://www.dotspatial.org)) component library. Supporting tools have also been developed to process and manipulate large climate data from different sources including the ENSEMBLES project (Haylock et al., 2008), the Future Flows project (Prudhomme et al., 2012) and the UKCP09 climate projections (Murphy et al., 2009b).

Key components of the modelling framework and supporting tools include:

- a GIS-based user-interface to manipulate and visualise spatial datasets such as hydrogeological data;
- functionality to support exploring aspects of hydrological model uncertainty including the sampling (random or Latin Hypercube Sampling – see Section 4.3.1.3) of hydrological parameter sets for two alternative hydrological model structures;
- functionality to enable alternative spatial resolutions in the application of climate data and the soil moisture accounting process to be considered;
- calculation and export of performance measures, river flow time-series and flow duration curves and other statistics;
- manipulation of climate data including extraction into specific temporal and spatial ‘slices’, calculation of catchment average values and re-projection between different co-ordinate systems.

### *3.3.1 Choice of hydrological model structures*

Recent studies into hydrological modelling uncertainty have suggested that multiple hydrological model structures should be considered (e.g. Wilby & Harris, 2006; Manning et al., 2009; Cloke et al., 2010; Christerson et al., 2012; Borgomeo et al., 2014) but there is little guidance made available to the water resources industry as to which model structures might be the most appropriate for strategic water resources planning in a given basin and why. Whichever model structures are used for a study it is important that their strengths and weaknesses (related to the context of the research) are assessed and understood prior to their subsequent use in prediction.

For this research, the CatchMod (also referred to as the TCM model structure) and the PDM-AQ model structures have both been integrated within the Kestrel - IHM modelling framework and examined in terms of their suitability for water resources planning in the Thames basin. The CatchMod model structure was selected because it is the model structure used by both Thames Water and the Environment Agency to model the Thames catchment and has been used in a significant volume of academic research relevant to this thesis (e.g. Wilby & Harris, 2006; Manning et al., 2009; Matrosov et al., 2013; Borgomeo et al., 2014; Walsh et al., 2015). The PDM-AQ model structure was selected because it is an established model structure also used in several academic and industry studies (e.g. Moore, 2007; UKWIR, 2007; Reynard et al., 2010; Prudhomme et al., 2012) and provides a conceptual contrast to the CatchMod structure in its formulation, particularly its use of parallel slow and fast pathways (similar to the majority of conceptual model



structures) and a probability distributed soil moisture store. Further details of each model structure are provided in sections 3.3.2 and 3.3.3 respectively.

In this thesis, each model structure has been run in a semi-lumped form for each study catchment, with the climatology averaged over each study catchment as input into the models. Each catchment was discretised into different hydrological response areas with the classifications based on the BGS 1:50,000 Solid and drift maps and informed by previous research of the Thames catchment (Bloomfield et al., 2009) which indicated a strong relationship between hydrogeological classification and the (base) flow response of the Thames.

Four distinct hydrogeological classes have been used, informed by Bloomfield et al. (2009), with aquifers divided into those where fracture flow is dominant (e.g. limestone aquifers) and aquifers where intergranular flow is dominant (e.g. consolidated sandstone aquifers). Aquitards were similarly split into two classes; non-permeable surficial deposits, and consolidated aquitards. Table 3-1 provides an overview of the modelled hydrogeological characteristics, along with the baseflow index (as reported on NRFA), for each of the sub-catchment areas examined in this research. The baseflow index is a measure of the proportion of the river runoff that derives from stored sources; the more permeable the rock, superficial deposits and soils in a catchment, the higher the baseflow and the more sustained the river's flow during periods of dry weather (Gustard et al., 1992).

To investigate the potential benefits that might be offered by making use of the gridded climatological inputs directly, and performing the soil moisture accounting on a 'gridded' basis, semi-distributed forms of the two model structures were also examined for the Thames catchment area draining to Kingston.

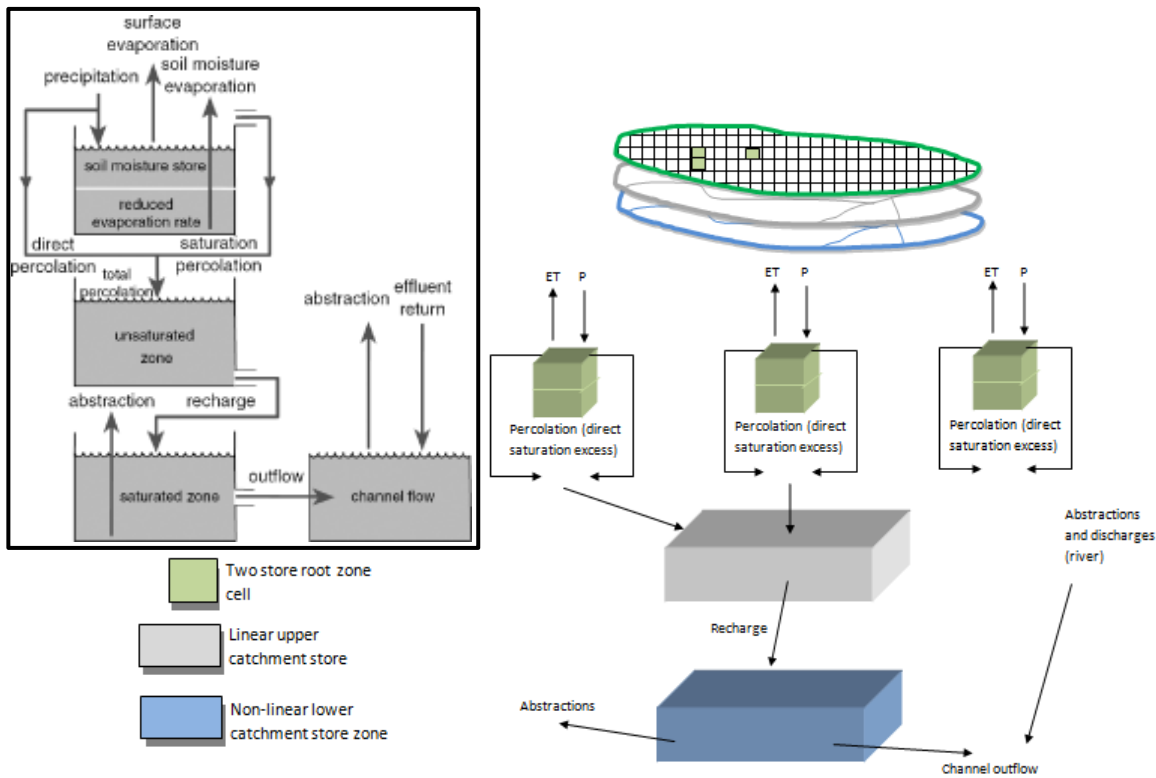
**Table 3-1: Hydrogeological classifications (after Bloomfield et al., 2009) for the modelled catchments.**

Catchment	Area (Km <sup>2</sup> )	Baseflow Index	Area of each hydrogeological classification			
			Fractured aquifer	Intergranular aquifer	Consolidated aquitards	Low permeability surficial deposits
<b>Thames@ Kingston (39001)</b>	9,969	0.63	38%	10%	30%	22%
<b>Windrush @ Newbridge (39006)</b>	370	0.86	91%	5%	2%	2%
<b>Wey @ Tilford (39011)</b>	380	0.72	7%	59%	0%	33%
<b>Kennet @ Theale (39016)</b>	1,008	0.88	54%	6%	11%	28%
<b>Cherwell @ Enslow Mill (39021)</b>	524	0.65	13%	0%	83%	4%

### 3.3.2 The Thames Catchment Model (TCM)

The Thames Catchment Model (TCM) is a conceptual rainfall-runoff model which includes a two-layer soil moisture accounting model which generates effective rainfall that is routed to the basin outlet through a linear and quadratic reservoir stores in series. A *Direct Percolation* pathway is available that allows a proportion of the rainfall that exceeds the PET rate to bypass the soil horizon during periods of soil moisture deficit to reflect responses in groundwater fed rivers – a detailed description of the TCM model structure, including the process equations, can be found in Wilby et al. (1994). An interesting feature of this model is the ability to develop infinitely (as in there is no upper limit on the magnitude of deficit that can be realised) large soil moisture deficits implemented to reflect catchment responses during the 1976 drought (Wilby et al., 1994).

Figure 3-2 presents a conceptual overview of the Thames Catchment Model (TCM) and how it has been applied in a semi-distributed (gridded) form as part of this research. This semi-distributed form processes the climatic inputs and soil moisture calculations (the two rootzone stores marked as “soil moisture store” and “reduce evaporation rate” store in the original CatchMod formulation, shown in the top-left of the figure) at a grid cell scale, with the cells classified into different response types for parameterisation. The routing of flows through the sequential linear upper catchment store (“unsaturated store” in the original CatchMod formulation) and non-linear lower catchment store (“saturated zone” in the original CatchMod formulation) is processed on a ‘semi-lumped’ basis as per the original semi-lumped form of TCM. There is no lateral flow between model elements (i.e. between grid cells).



**Figure 3-2: The TCM conceptual model structure (semi-distributed form shown) with original CatchMod formulation (reproduced from Cloke et al., 2010) shown in top-left box for reference.**

### 3.3.3 The Probability Distributed Model (PDM)

The Probability Distributed Model (PDM) is described as a fairly general conceptual rainfall-runoff model (Moore, 2007) which has been employed in many forms to suit different modelling objectives (e.g. Bell et al., 2007; Reynard et al., 2010; UKWIR, 2007).

All forms of PDM include a 'mass-balance' probability distributed soil moisture accounting component, with resulting direct runoff and recharge routed via 'slow' and 'fast' pathways to the basin outlet. A Pareto distribution is typically used (and has been used in this research) to describe the distribution of the storage capacity across a catchment, with the distribution shape altered to reflect different proportions of deep or shallow stores. If the storage capacity at a point is exceeded, direct runoff occurs, otherwise water remains in storage with losses to evaporation and via recharge to the groundwater store. Both the direct runoff and the recharge component use a storage element to represent surface storage ('fast') and groundwater storage ('slow') and act as a delay in the system to represent different catchment characteristics. The catchment river flow output combines the discharges from the surface and groundwater stores (see Moore, 2007, for a detailed description, including the process equations, of the PDM model structure).

For this research, a form of the PDM model structure appropriate for use in permeable, groundwater dominated catchments has been used based on the findings of Moore (2002). This

adopts a quadratic non-linear store (as used in the TCM model structure) for the ‘slow’ (groundwater) flow pathway considered appropriate for unconfined aquifers (Moore, 2002). A *Soil Tension* parameter is included to reflect the soil moisture held under tension and beyond which drainage does not occur but from which moisture is still available for evaporation. For the PDM-AQ model structure, in areas hydrogeologically (Bloomfield et al., 2009) classified as either ‘aquitard’ or ‘low permeability surficial deposits’, a slow tank is not used in these areas as a significant aquifer is not considered to be present. In the remainder of the thesis, the PDM-AQ notation is used to emphasise that a specific form of the PDM model structure appropriate for groundwater dominated catchments has been adopted in this thesis.

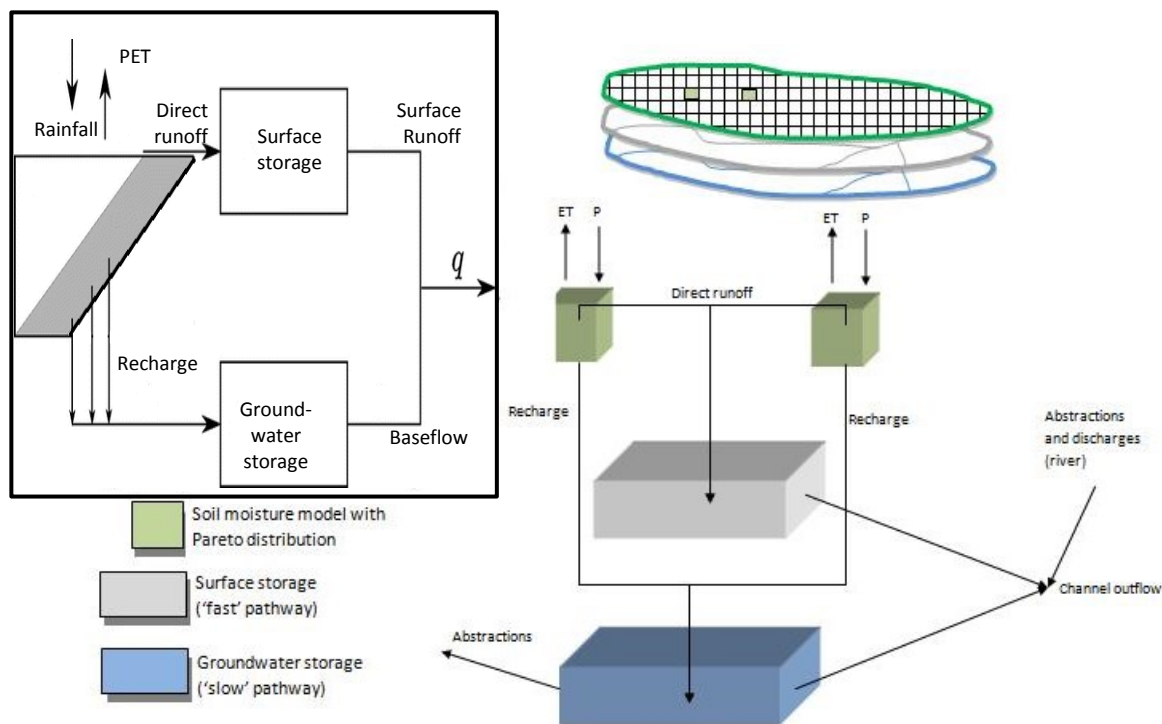
Figure 3-3 provides an overview of the PDM model structure and how a semi-distributed form has been applied in this thesis. Climatic inputs and soil moisture calculations are processed for each grid cell using the original PDM model formulation, with each cell classified into different response types for parameterisation as described below. The routing of direct runoff and recharge through the parallel ‘fast’ and ‘slow’ pathways respectively (through the “surface storage” and “groundwater storage” units in the original PDM model formulation) are processed on a ‘semi-lumped’ basis with no lateral flow between model elements (i.e. between grid cells). These subsequently generate the surface water and baseflow components of the outflow.

### 3.4 Data used to drive the hydrological models

This section describes the data used to drive the hydrological models which are the river flow records used to assess model performance and the precipitation and PET data used as inputs to the models.

#### 3.4.1 *Historical river flow data*

For Kingston, the ‘natural’ flow record used by Thames Water in their water resources and drought planning was used in this research. The ‘natural’ flows (distinct from a fully ‘naturalised’ record) for the Thames at Kingston are the gauged flows plus the recorded abstractions made by Thames Water and Affinity Water (Thames Water, 2013) just upstream of Kingston. This flow series is considered to represent the flow available at Thames Water’s abstractions close to Kingston. For the other study catchments, the gauged flow record available from the National River Flows Archive (NRFA) has been used (refer to Table 3-2 for the periods used).



**Figure 3-3: The PDM-AQ conceptual model structure (semi-distributed form shown) with original PDM formulation (adapted from Moore, 2007) shown in top-left box for reference.**

In this research, each historical flow record was split into five discrete periods (similar to the approach adopted by Bárdossy et al., 2016) to evaluate model performance. Adopting multiple time periods (and sites) may help identify potential issues in the historical data or highlight how changes in the catchment might have altered the overall response noting that modelling studies usually assume that model parameters are constant over time. Further, using multiple periods allows each period to be potentially used as a validation period prior to being incorporated into the model conditioning dataset (e.g. see Beven & Binley, 2013).

Table 3-2 highlights these discrete periods along with the historical lowest Mean Summer Flow (MSF) between April and September (as used by Thames Water (2010) as an indicator of potential drought in London) in each period for each of the five study catchments. The records were split such that the dates used to divide the record did not coincide with periods of significantly low flow (refer to the analysis of the Standardised Precipitation Index in Section 3.4.2.3).

**Table 3-2: Lowest historical mean summer flow (April to September, m<sup>3</sup>/s) for five separate time-periods for five locations within the Thames basin.**

<b>Catchment / time period</b>	<b>1961 - 1970</b>	<b>1971 – 1980</b>	<b>1981 - 1992</b>	<b>1993 – 2000</b>	<b>2001 - 2008</b>
Thames@ Kingston (39001)	38.0 (1965)	18.1 (1976)	31.9 (1992)	27.8 (1997)	36.3 (2005)
Windrush @ Newbridge (39006)	1.8 (1965)	0.4 (1976)	1.2 (1990)	1.2 (1997)	1.6 (2003)
Wey @ Tilford (39011)	2.1 (1965)	1.3 (1976)	1.5 (1992)	1.4 (1997)	1.5 (2005)
Kennet @ Theale (39016)	4.5 (1965)*	2.3 (1976)	5.1 (1992)	3.9 (1997)	5.2 (2005)
Cherwell @ Enslow Mill (39021)	2.1 (1970)**	0.3 (1976)	1.2 (1990)	1.1 (1997)	1.3 (2003)

\* record starts in 1962

\*\* record starts in 1967

Table 3-3 reports nine selected flow percentiles (used in later chapters to assess model performance and the impacts of climate change) from the historical flow record for Kingston (39001) across the same five time-periods and demonstrates the relative dryness of the 1971 to 1980 period and the relative wetness of the 2001 to 2008 and 1961 to 1970 periods. In terms of trends, there is some evidence for increasing high flows (Q30 and above) but no trends evident with regards to medium to low flows. These contrasting periods are revisited later in this thesis.

**Table 3-3: Selected flow percentiles from the historical flow record at Kingston for five separate time-periods (calendar years).**

<b>Flow percentile</b>	<b>1961 - 1970</b>	<b>1971 – 1980</b>	<b>1981 - 1992</b>	<b>1993 – 2000</b>	<b>2001 - 2008</b>
<b>Q1</b>	299.0	318.5	309.5	355.5	339.0
<b>Q5</b>	217.0	224.8	208.0	250.4	254.0
<b>Q10</b>	168.0	169.0	157.8	192.0	189.2
<b>Q30</b>	89.2	90.6	88.9	92.5	95.9
<b>Q50</b>	64.9	55.2	57.5	59.6	63.1
<b>Q70</b>	42.8	38.2	38.6	37.6	47.8
<b>Q90</b>	30.3	27.3	26.7	25.3	28.6
<b>Q95</b>	27.2	23.9	22.1	21.6	24.6
<b>Q99</b>	24.0	11.8	18.6	18.4	20.4

### 3.4.2 Historical climate data

Historical climate data were taken from the Ensembles E-Obs version 3.0 (Haylock et al., 2008) and the UK Met Office (2013) historical gridded datasets as summarised in Table 3-4.

**Table 3-4: Sources of historical climate data used in this thesis.**

<b>Source</b>	<b>Weather variables</b>
E-Obs (Haylock et al., 2008)	Daily precipitation and mean daily temperature on a 0.25 degree regular latitude-longitude grid from 1950 to 2009. Version 3 used in this thesis (refer to <a href="https://www.ecad.eu/download/ensembles/oldversions.php">https://www.ecad.eu/download/ensembles/oldversions.php</a> )
Met Office gridded data (Met Office, 2013)	Monthly records of mean daily temperature (1950 to 2009), wind speed (1969-2009), relative humidity, cloud cover, sea level pressure, vapour pressure (all 1961 – 2009) all on a 5km grid

During the course of this research, CEH has published new gridded precipitation (GEAR - Tanguy et al., 2016) and PET datasets (CHESS - Robinson et al., 2016) but these have not been used in the impacts modelling presented in this thesis. These new data could provide a useful comparison of the data used here albeit that these are likely to be underpinned by similar weather station data. For PET, Robinson et al. (2016) made use of the Penman-Monteith equation (Allen et al., 1998) and a modified form of this equation has also been used in this research as described in the following section.

#### 3.4.2.1 Estimation of potential evapotranspiration (PET)

This research considers two alternative methods; an empirical temperature-based method (Oudin et al., 2005) and a modified form of the Penman-Monteith equation (Allen et al., 1998, Prudhomme & Williamson, 2013). These two methods were also used because they can exploit the variables available in the UKCP09 climate change projections, albeit assuming no change in

mean wind speed for the modified Penman-Monteith equation (Prudhomme & Williamson, 2013), which is discussed further in Chapter 5.

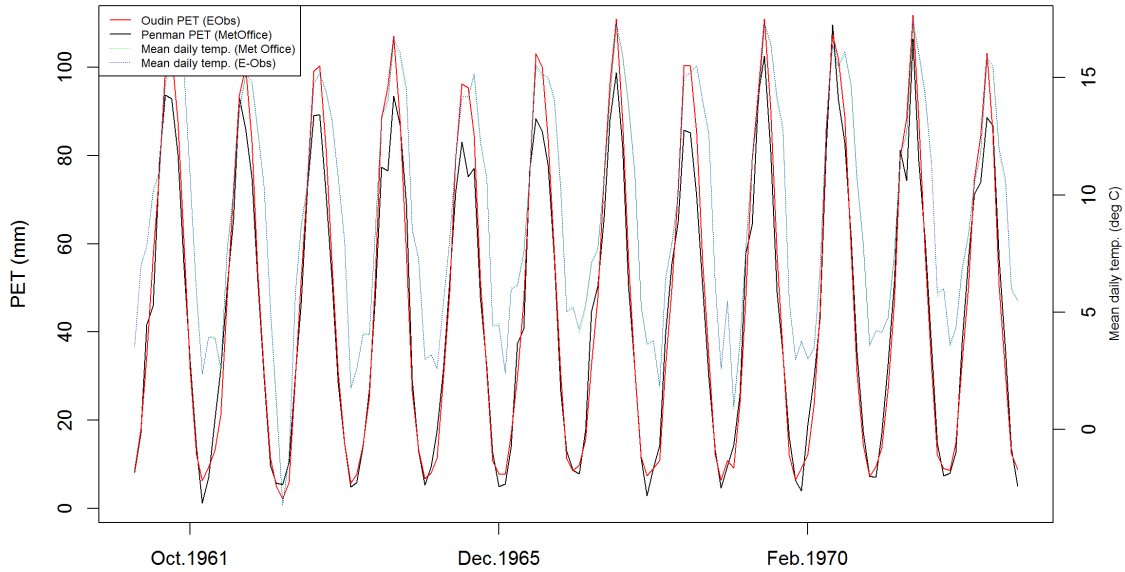
To derive daily values of PET using the temperature-based Oudin method, mean daily temperature values from the Ensembles historical dataset (Haylock et al., 2008) were used as input. For the modified form of the Penman-Monteith equation, the Met Office (2013) monthly gridded data of temperature, wind speed, cloud cover, sea level pressure and vapour pressure were all used with the resulting monthly estimates then downscaled to daily values by scaling the daily values calculated under the Oudin method by the difference in monthly totals.

Figure 3-4 (a to d) plot the monthly estimates of PET using the two methods across the Thames catchment from 1961 to 2008 along with the recorded values of mean monthly temperature from both the Ensembles and Met Office datasets. The two PET methods exhibit similar profiles but differ as to which produces the higher rates of PET during the summer periods. Typically, summer values calculated using the Oudin method peak at higher monthly rates, including for the particularly dry periods (identified later in this chapter) such as the summers of 1965, 1991, 1992, 1997, 2003 and 2005. However, the opposite is true during the hot, dry summer of 1976 and also for the summer of 1990.

During the 1961 to 2008 time-period the average annual PET calculated using the modified Penman-Monteith method is 636mm compared to 590mm using the Oudin method. From May to September the totals are 443mm and 434mm respectively, and from October to April, 193mm and 156mm hence the estimated rates of PET during winter months are much higher under the modified Penman-Monteith method than under the Oudin method. This subsequently affects the rates of winter groundwater recharge and subsequently flows during the following summer, although this may be mitigated by the hydrological calibration process. This is discussed further in Chapter 4.

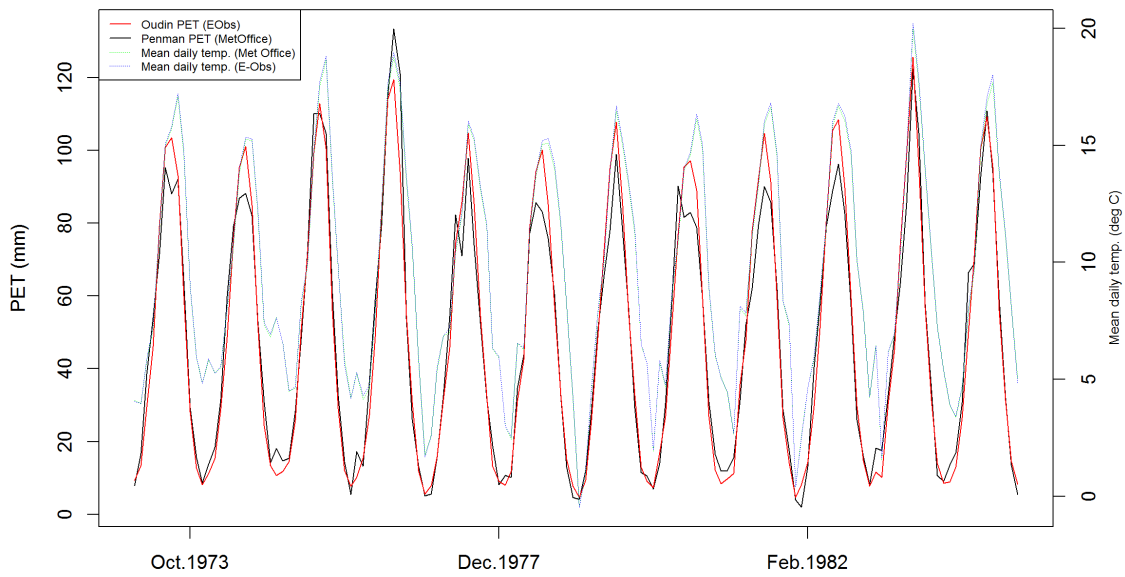


### Oudin and Penman Monteith (modified) PET



D:\Work\PhD\FF\_Temp\_Analysis\Penman\_PET\_Baseline\_Comparison / 29 Oct 2014 / cjc

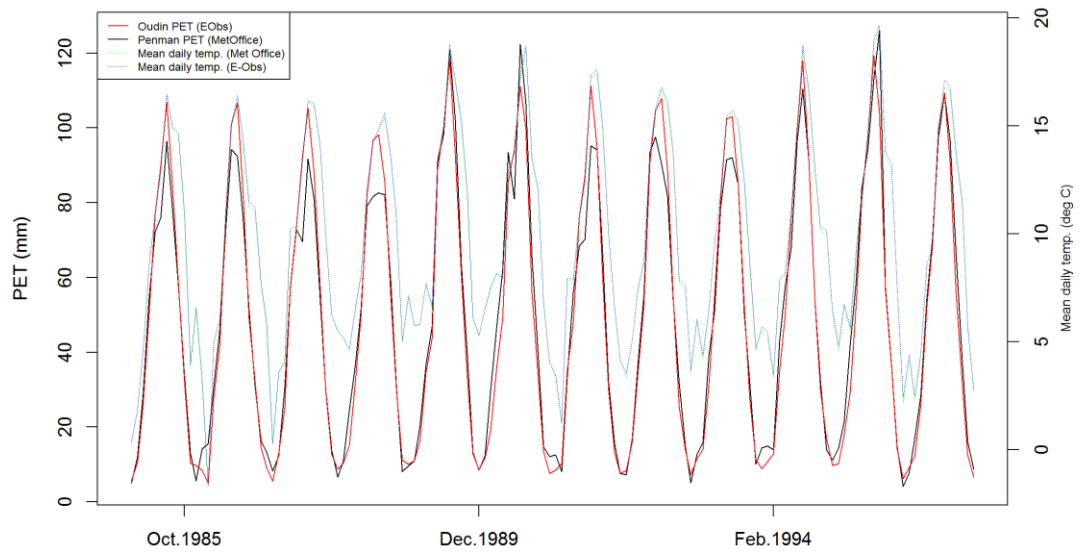
### Oudin and Penman Monteith (modified) PET



**Figure 3-4a and b: Comparison of estimated monthly total PET using the Oudin equation (red) and the modified form of the Penman-Monteith equation (black) from 1961 to 1972 (Figure 3-4a, upper plot) and 1973 to 1985 (Figure 3-4b, lower plot) across the Thames basin to Kingston<sup>1</sup>.**

<sup>1</sup> Note that Met Office and E-Obs mean daily temperature are very similar and hence difficult to distinguish in the plot.

### Oudin and Penman Monteith (modified) PET



### Oudin and Penman Monteith (modified) PET

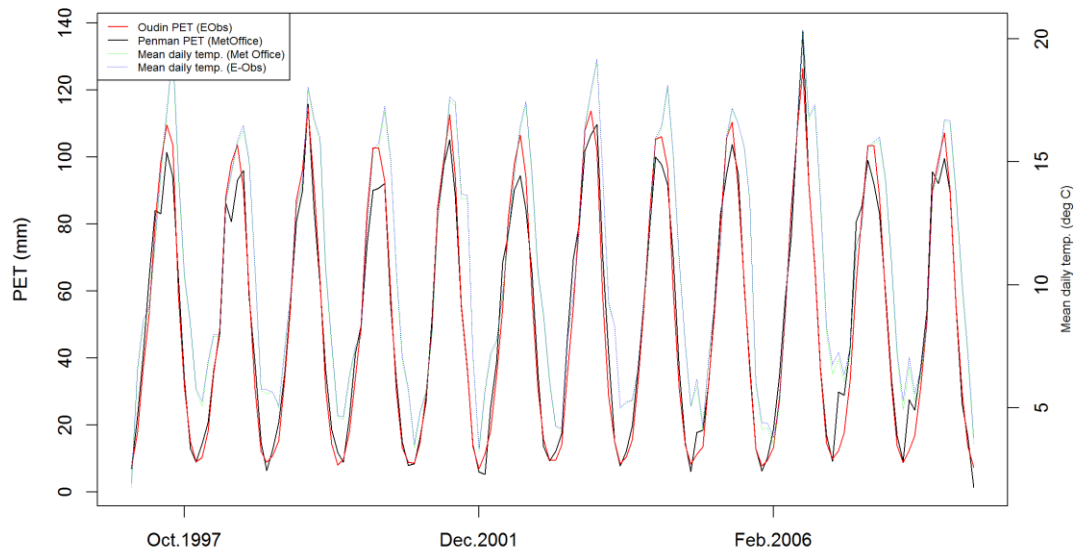
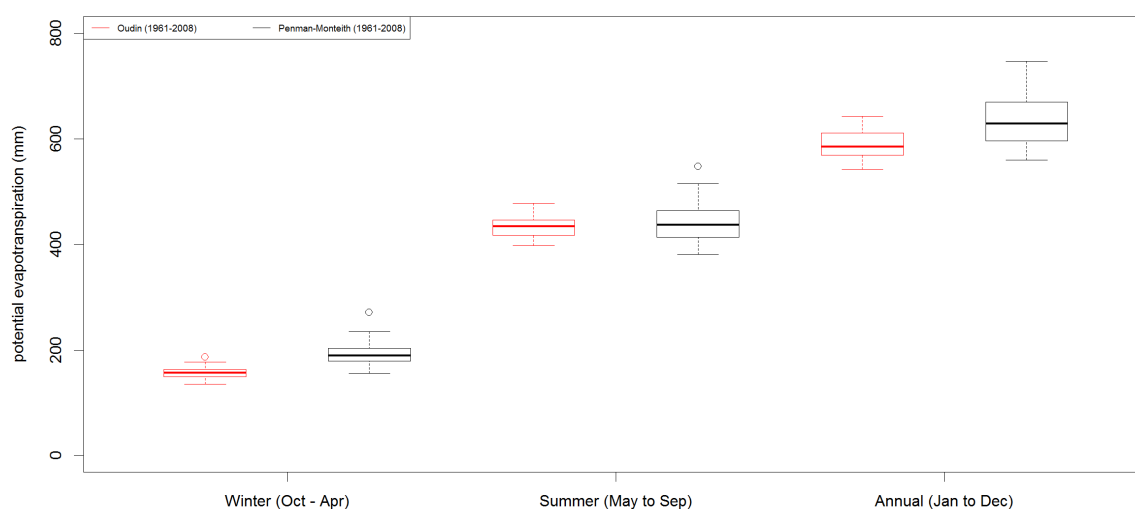


Figure 3-4c and d: Comparison of estimated monthly total PET using the Oudin equation (red) and the modified form of the Penman-Monteith equation (black) from 1985 to 1996 (Figure 3-4, upper plot) and 1997 to 2008 (Figure 3-4d, lower plot) across the Thames basin to Kingston.<sup>2</sup>

<sup>2</sup> Note that Met Office and E-Obs mean daily temperature are very similar and hence difficult to distinguish in the plot.

The distribution of these values is presented in Figure 3-5 which further highlights that summer PET rates are broadly similar but that rates during winter are much lower using the Oudin method and this is the predominant cause of the differences in mean annual totals between the two methods. This is consistent with the findings of Prudhomme & Williamson (2013) who reported that, in general, winter estimates of PET by radiation and temperature based equations, including the Oudin method, are lower than for Penman-Monteith based methods. Prudhomme & Williamson (2013) also demonstrated significant differences between the two methods used in this research during the winter months of November through to February (based on the 1961 – 1990 time-slice).



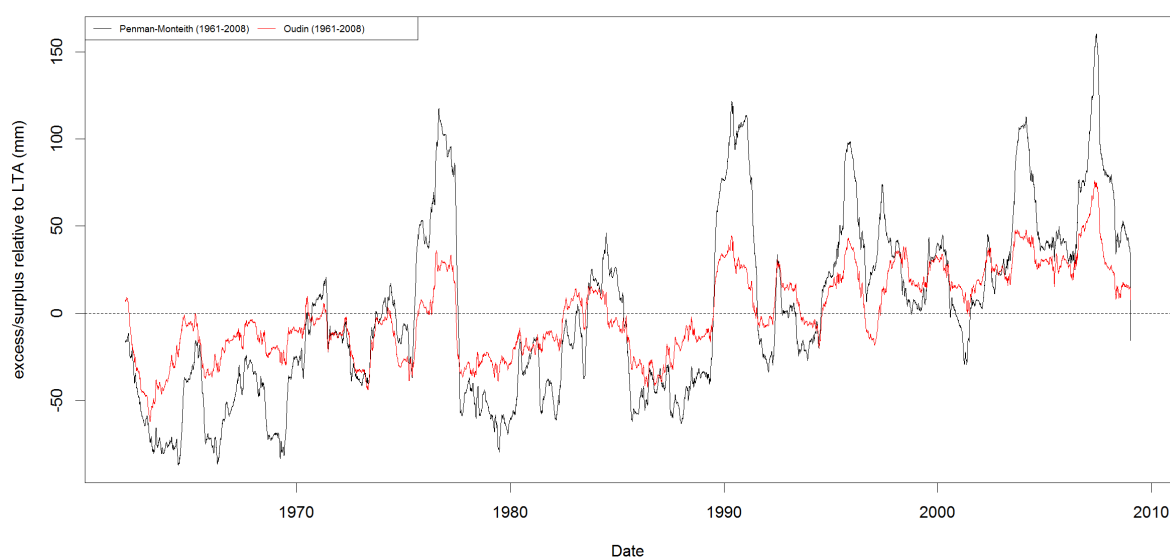
**Figure 3-5: Boxplot showing the distribution of winter (October to April), summer (April to September) and annual (January to December) rates of PET (to Kingston) calculated using the Oudin (red) and Penman-Monteith methods (black) between 1961 and 2008.**

For water resources planning, the cumulative balance between precipitation and PET is of interest in considering water availability - particular in groundwater dominated catchments (Charlton & Arnell, 2014) such as the Thames basin. However, previous water resources studies (e.g. Manning et al., 2009; Christerson et al., 2012; Borgomeo et al., 2014) have seldom provided a detailed analysis of the PET records during periods of low flow. It should be emphasised that it is actual evapotranspiration, which often occurs at less than the potential rate due to crop stress arising from limited soil moisture during low flow periods (e.g. Rushton et al., 2006; Zhao et al., 2013), that impacts water availability.

The potential significance of PET can be seen in Figure 3-6 which presents the historical PET record estimated using both methods, in terms of a rolling 365-day cumulative surplus/deficit relative to their respective long-term average (LTA) 365-day cumulative PET (1961 to 2008). Figure 3-6 highlights a number of interesting features. Firstly, the relatively low flow periods (see Section

3.4) of 1976 and 1990/1991 coincided with relatively high cumulative values of PET. Secondly, both methods used to estimate PET suggest a general tendency (trend) for increasing rates of PET through the 1961 to 2008 period, consistent with the general warming across the UK reported by others (e.g. Jenkins et al., 2009).

Finally, it can be seen that the modified Penman-Monteith method exhibits greater temporal variability, generating larger cumulative deficits and surpluses over the historical record. This is a consequence of employing a more physically based model rather than an empirically derived relationship. The more physically based method allows for the interplay between different weather variables (e.g. relative humidity) to enhance or dampen the rate of PET for a given temperature compared to the empirical based approach which implicitly ties the PET rate to temperature alone. This suggests that a more physically-based model might be more appropriate when considering more extreme weather sequences, such as droughts which drive the water resources planning process, to ensure such variability is captured in the modelling. However, methods such as Oudin have much lower data requirements which may be considered an advantage when outputs from climate models are to be used, particularly as the confidence associated with some of the variables required for more physically-based models may be lower (Prudhomme & Williamson, 2013). This is discussed further in Chapter 5.

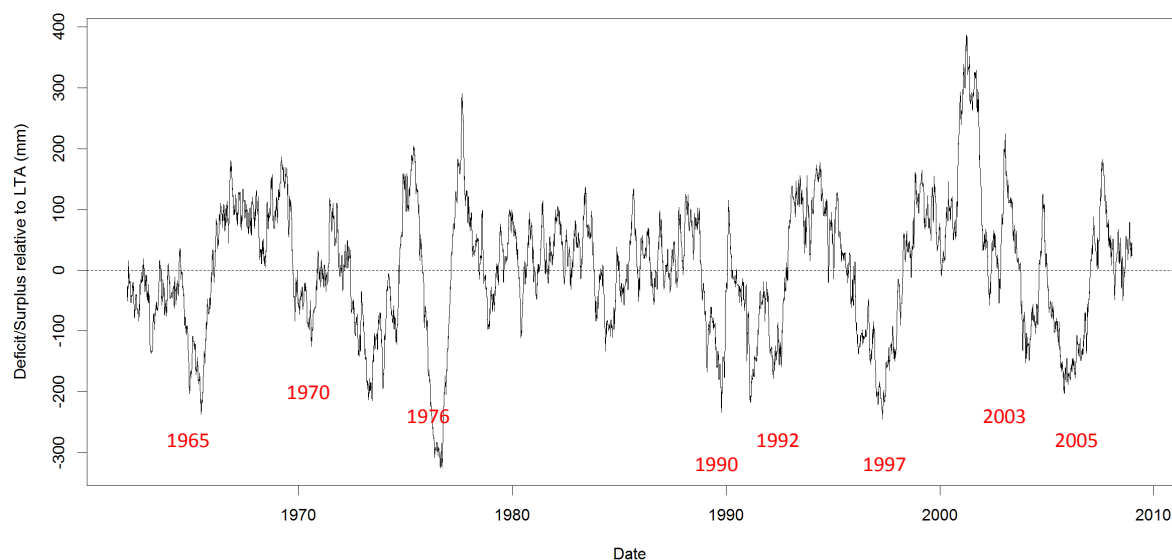


**Figure 3-6: Rolling 365-day cumulative deficit/surplus PET using both the Oudin (red) and modified Penman-Monteith (black) methods relative to their respective historical long-term average (rolling 365-day catchment average over the Thames catchment to Kingston).**

#### 3.4.2.2 Historical precipitation record

Figure 3-7 presents the historical precipitation record in terms of a rolling 365-day surplus/deficit from the long-term average (LTA) 365-day precipitation (1961 to 2008) and highlights the low flow periods reported in Table 3-2.

Table 3-5 reports summary information including the mean precipitation, the number of years within each period (refer to Table 3-2) when the rolling deficit, relative to the long term average (1961 to 2008), exceeds 100mm (equivalent to a reduction of 15% of the 1961 to 2008 average) along with the maximum 365-day precipitation deficit within each period. The deficit threshold of 15% has been informed by the identification of notably dry periods by Marsh et al. (2014) but it is recognised that this is subjective. The relative significance of these ‘dry’ years is examined further in section 3.4.2.3 and is referred to again during the hydrological modelling and water resources modelling presented in chapters 4 and 6.



**Figure 3-7: Rolling 365-day cumulative deficit/surplus precipitation relative to the historical long-term average (rolling 365-day catchment average precipitation over the Thames catchment to Kingston).**

From Figure 3-7 there appears to be no obvious trend in mean annual precipitation through the 1961 to 2008 time-period in contrast to the evidence shown in Figure 3-6 regarding PET. This is consistent with the findings of Jenkins et al., (2009) that there has been little change in mean annual precipitation over the last 50 years.

**Table 3-5: Comparison of mean annual precipitation and the frequency and severity of dry periods in the Thames catchment to Kingston for different periods in the historical record.**

Period	Mean 365-day Precipitation (mm)	No. of years where max deficit to LTA > 100mm	Maximum deficit (mm) and year)
1961 - 1970	673	3	238 (1965)
1971 – 1980	658	4	325 (1976)
1981 - 1992	667	4	234 (1989)
1993 – 2000	712	1	245 (1997)
2001 - 2008	668	2	203 (2005)
1961 - 2008	674	14	325 (1976)

#### 3.4.2.3 Standardised Precipitation Index (SPI) and Standardised Precipitation and Evapotranspiration Index (SPEI)

The Standardized Precipitation Index (SPI) is designed to quantify the precipitation deficit for multiple timescales to reflect the impact of drought on the availability of the different water resources (McKee et al., 1993). The World Meteorological Organization (WMO) has adopted the SPI for use by national meteorological and hydrological services worldwide to characterize meteorological droughts (WMO, 2012).

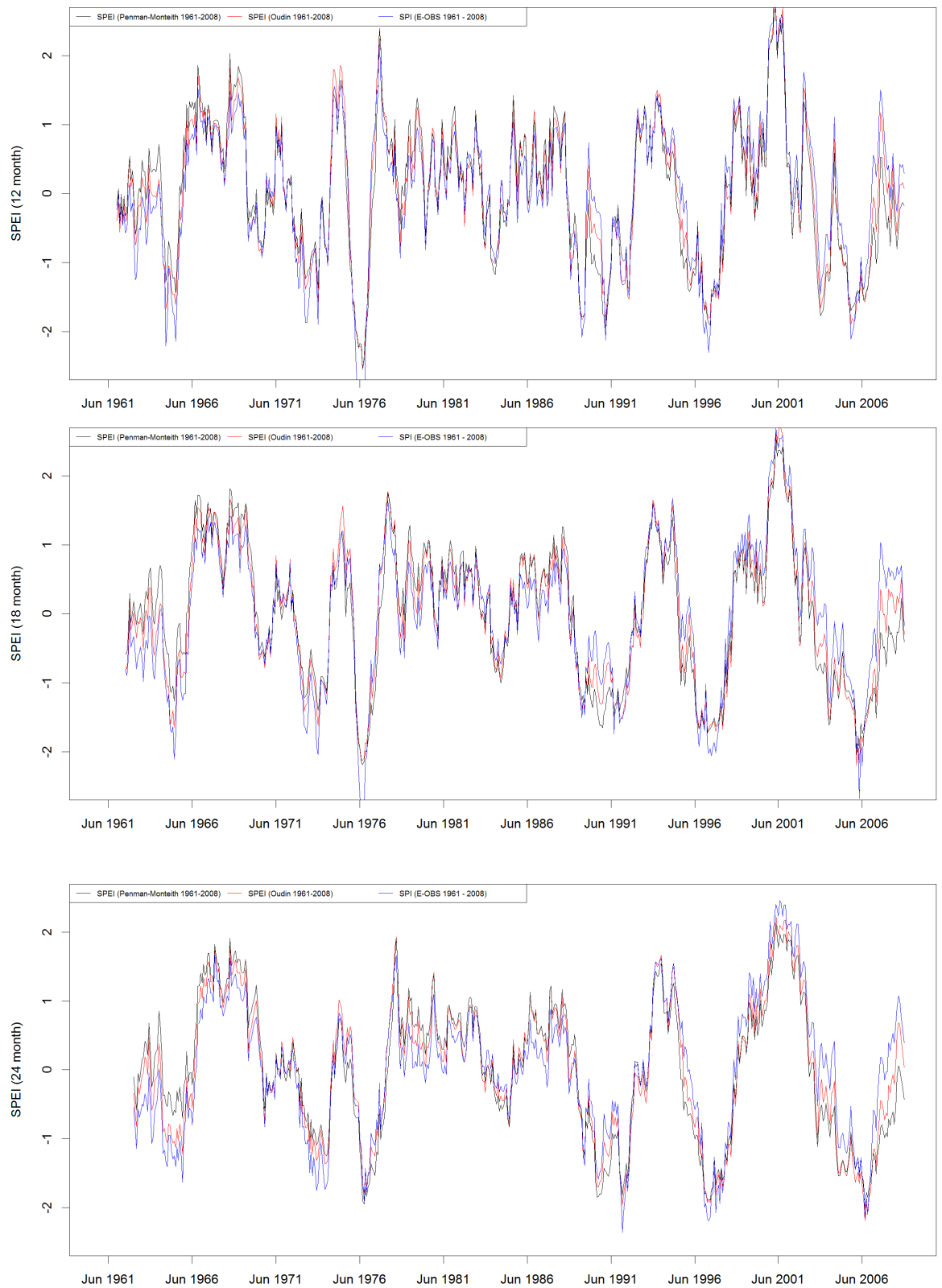
The SPI calculation for any location is based on the long-term precipitation record for a desired period and this record is fitted to a probability distribution which is then transformed into a normal distribution so that the mean SPI for the location and desired period is zero (World Meteorological Organization, 2012). A gamma distribution, found to provide the best model for describing monthly precipitation over most of Europe (Lloyd-Hughes & Saunders, 2002; Stagge et al., 2015) has been used in this research. Positive SPI values indicate greater than median precipitation and negative values indicate less than median precipitation. As a result of the transformation to a normal distribution, McKee et al. (1993) reports that SPI values between -1.5 and -1.99 are considered severely dry and values less than -2 as extremely dry and that these would occur with return periods of approximately 1 in 20 and 1 in 50 years respectively. It should, however, be noted that these are being calculated using a record length of 48 years and thus there may be significant statistical uncertainty in associating return periods of such magnitude.

However, the SPI relies on precipitation only and the Standardized Precipitation Evapotranspiration Index (SPEI) was developed to also take into account both precipitation and PET in determining drought (Vicente-Serrano et al., 2010). As an index it has been shown to be useful in characterising the evolution of drought using climatological variables only (Beguería et al. 2014). For this research, the log-logistic distribution was adopted (as recommended by Vicente-Serrano et al., 2010; 2016 and used by Tanguy et al., 2015) although it is recognised that the choice of appropriate probability distribution has recently been subject to debate in the academic literature (see Stagge et al., 2015; Vicente-Serrano & Beguería, 2016). The choice of probability distribution influences the resulting SPEI values, particularly at the extremes, which has implications for comparing drought severity across different sites. However, the choice of probability distribution has no implications for the remainder of the research presented in this thesis so has therefore not been considered further.

Figure 3-8 (a to c) present the 12-month, 18-month and 24-month SPEI and SPIs for the Thames catchment to Kingston. For estimating SPEI, both values using Oudin and the modified Penman-

Monteith methods for estimating PET are shown, noting that these SPEI values are *standardised* indices. The 12, 18 and 24 month accumulation time-periods have been chosen because the Thames catchment is groundwater dominated and because the London water resource system has significant storage available (fully replenished surface water storage alone could provide over 100 days of water demand in London – based on Birks et al., 2001 and values presented in Thames Water’s latest WRMP, 2014) with which to mitigate short, sharp drought sequences.

Whilst the periods identified as being more ‘droughty’ (e.g. SPI values less than -1.5) are consistent with the results presented in Figure 3-7, Figure 3-8 highlights that different periods appear relatively more significant than others as the time-window over which the indices are calculated is altered. For example, the relative intensity of the 1976 and 2006 events change significantly when using the 12-month and 24-month accumulation time-windows. Both the 24-month SPEI and SPI values highlight how the 2003 and 2005 events, and similarly the 1990 to 1992 period, combine to form single, more sustained droughts. In contrast, the 1976 drought is a much shorter but more intense drought using these indices and over the longer time-windows its relative severity reduces.



**Figure 3-8a to c: SPEI (black for modified Penman-Monteith PET method and red for Oudin PET method) and SPI (blue) indices for 12-month (a, top plot), 18-month (b, middle plot) and 24-month (c, lower plot) time windows across the Thames catchment to Kingston.**



The two methods for estimating PET are broadly consistent in their SPEI values, noting that these are standardised indices and this is consistent with the findings of Beguería et al. (2014) and Stagge & Tallaksen (2014) which found relatively lower differences due to using different equations to estimate PET, particularly in relatively humid areas. The general profile of SPI index is also consistent with the SPEI indices although the SPI values demonstrate high values at the extremes which vary depending upon the time-window used. This may be a consequence of the relative balance between precipitation and PET during the different seasons and the relative contribution of different seasons within the calculated values (e.g. an 18-month metric will include two summers and one winter and vice-versa during its calculation).

In contrast to the increasing trend in PET observed in Figure 3-6, a similar trend is not apparent in SPEI. This, in part, may be due to this relatively drought poor period between 1976 and 1990 and the different temporal characteristics of the notably drought events either side of this period. These plots also highlight that the different time-periods considered in the modelling, apart from this relatively drought poor period, each contain one of the more notably dry events which are discussed further in chapters 4 and 6.

### 3.5 Chapter summary

This chapter has introduced the case study catchments that are the focus of this research along with a description of the hydrological modelling framework and the two conceptual model structures considered in this research, the TCM and PDM-AQ model structures, and their application to the study catchments. The historical climate and flow data used to drive and condition these models have been examined including a comparison of two contrasting methods for estimating PET; an empirical temperature-based method (Oudin et al., 2005) and a modified form of the Penman-Monteith equation (Allen et al., 1998; Prudhomme & Williamson, 2013).

This chapter has demonstrated that both methods used to estimate PET suggest a general tendency (trend) for increasing rates of PET through the 1961 to 2008 period, consistent with the general warming across the UK reported by others (e.g. Jenkins et al., 2009). No obvious trend is evident within the precipitation record. Typically, summer PET values calculated using the Oudin method peak at higher monthly rates, including for the particularly dry periods (identified later in this chapter) such as the summers of 1965, 1991, 1992, 1997, 2003 and 2005. However, the opposite is true during the hot, dry summer of 1976 and also for the summer of 1990. Annual estimates are higher using the modified Penman-Monteith method, primarily related to estimates during winter months. Further, the modified Penman-Monteith method is shown to exhibit

greater temporal variability, generating larger cumulative deficits and surpluses relative to their respective long-term average 365-day cumulative PET values than the Oudin method.

SPEI (Vicente-Serrano et al., 2010) and SPI (McKee et al., 1993) indices using 12, 18 and 24 month accumulation time-periods have been used to examine the presence and relative severity of droughts in the historical climate records. Both the 24-month SPEI and SPI values highlight how the 2003 and 2005 events, and similarly the 1990 to 1992 period, combine to form single, more sustained droughts. In contrast, the 1976 drought is a much shorter but more intense drought using these indices and over the longer time-windows (18 and 24 month accumulation periods) its relative severity reduces.

Chapter 4 describes the development of a 'multi-site' hydrological model ensemble to provide spatially coherent flows across the study locations in the Thames catchment. The skill of this model ensemble is also reviewed in Chapter 4 with a particular emphasis on performance measures relevant to water resources planning prior to its application in assessing climate change impacts (Chapter 5) and the reliability of London's public water supplies (Chapter 6) later in this thesis.



## Chapter 4 A hydrological model ensemble for water resources planning under uncertainty

### 4.1 Introduction

For complex water resources systems such as London's public water supply managed by Thames Water (2014), water resource system models are used to assess the available resources and plan how drought events will be managed. These water resource system models require river flow series as input typically generated using hydrological models driven by time-series of precipitation and PET. These hydrological models are also used to consider the potential impacts of climate change on future resource availability and to assess the potential yield from alternative surface water resource options (e.g. a new river intake or storage reservoir) that could be implemented to increase the available resource in the future (Thames Water, 2014). Hydrological models are also used to support water companies in their drought planning to estimate the potential risk (considering both likelihood and consequences) of developing critical water shortages, and to define the associated 'triggers' that signify when drought management options should be introduced in order to maintain supplies (e.g. Thames Water, 2010).

### 4.2 Objectives for the research presented in this chapter

The aim of the research is to complement and extend the existing knowledge on hydrological modelling uncertainty in the context of the water resources planning process presented in the literature review (see Section 2.3.3). The research questions (see Section 1.2) that are specifically posed for Chapter 4 are:

- RQ1.** How important are the uncertainties associated with hydrological model structures, model parameterisation and the method used to estimate PET and what are their influences on the flow characteristics of relevance to the water resources planning process?
- RQ4.** How do different hydrological models differ in their reproduction of river flows and the specific metrics of interest to water resources planners across the Thames catchment and how does this reproduction vary over different time-periods?

### 4.3 Development of a Multi-Site Model Ensemble (MSME) for water resources planning

Chapter 3 introduced the locations of interest (see Figure 3-1); the two hydrological model structures used (see Section 3.3); the input data and two alternative methods for estimating PET (see Section 3.4). This chapter describes the development of a hydrological model ensemble to provide spatially coherent river flow predictions at multiple sites of interest that can be accompanied with an assessment of the relative significance of different sources of hydrological modelling uncertainty.

The second part of this chapter examines how the model ensemble reproduces key flow characteristics of interest to the water resources planner along with the sensitivity of water resource relevant flow metrics to the uncertainty associated with hydrological model parameterisation, model structures and the two methods used to estimate PET (see Section 3.4.2.1) prior to the ensemble being taken forward to assess the impacts of climate change (Chapter 5) and the performance of London's water supply system (Chapter 6).

#### 4.3.1 *Uncertainty evaluation*

Traditional hydrological modelling studies typically split the historical record into separate calibration / validation data sets in which one period of observations is used in calibration and another separate period is used to check that the model predictions are satisfactory (e.g. Beven, 2001). A variation of this is to calibrate the models on the wetter parts of the historical record and validate on the drier half as a proxy for the anticipated impacts of climate change (e.g. Vaze et al., 2010; Oni et al., 2016) although, as noted by Brigode et al. (2013), it seems difficult to provide general guidelines for calibrating hydrological models for climate change studies.

For this research, a Generalised Likelihood Uncertainty Estimation (GLUE) methodology (Beven & Binley, 1992) has been adopted based on all five periods (see section 3.4.1) used in the model conditioning. Alternative approaches to quantifying uncertainty are discussed in Section 2.1.3. The implications of adopting a GLUE approach to characterising and quantifying uncertainty are recognised and other approaches, such as the use of formal Bayesian statistics (e.g. Kavetski & Fenicia, 2011; Euser et al., 2013) promoted by Clark et al. (2012) or the extended GLUE approach (e.g. Liu et al., 2009) described in Section 2.1.3 are alternatives. However, these alternative and extended approaches are subject to their own limitations and evidence requirements as described in Section 2.1.3. The transparent and reproducible approach adopted here is a considered a significant but practicable step beyond current water resources industry practice to provide a mechanism by which to evaluate the validity, strengths and weaknesses of alternative

hydrological model structures and explore the potential role that different sources of uncertainty play in the water resources planning process.

As highlighted in Chapter 2, previous research that has explored the potential significance of different sources of uncertainty (hydrological and climate change) in relation to water resources planning have typically used one or two performance metrics at a single location in each catchment, often considering such measures over a single time period, to identify ‘behavioural’ models subsequently retained for use in prediction (e.g. Manning et al., 2009; Cloke et al., 2010; Charlton & Arnell, 2014; Borgomeo et al., 2014). In the development of the hydrological model ensemble for this research, multiple locations (see Section 3.2), multiple time-periods (see Section 3.4) and multiple performance criteria (as described below) have been used.

#### 4.3.1.1 Performance measures

The six measures chosen for this research have been informed by this previous research and the objectives of the overall research. These performance measures are set-out in Table 4-1.

The performance measures include one that is directly relevant to the reproduction of sustained low flow sequences and the triggering of strategic supply options (*Drought\_error*) along with measures associated with overall model performance across all flow conditions (*Volume\_Error*, *NSE*, *FDC\_full*). The emphasis on low flows has been further emphasised by including *LogNSE* (NSE on the natural log of flows) and *FDC\_low* which focuses on the reproduction of the flow duration curve at the lower end. Adopting a range of performance metrics, locations and periods also potentially enables subsequent analysis of those aspects of the historical record which the model structures struggle to reproduce well and for which model structure improvements or further examination of the historical record may be warranted (Beven & Binley, 2013).

To provide the assessment of uncertainty, the six measures described above are subsequently used to reject or weight models under a Generalised Likelihood Uncertainty Estimation (GLUE) methodology (Beven & Binley, 1992 – see Section 2.1.3).

#### 4.3.1.2 Description of the GLUE methodology adopted for this research

In this research, the six water resources relevant performance measures, for each of the simulation periods were combined using a multiplicative formulae (see Beven & Binley, 1992; Beven & Freer, 2001b for details) following a Bayesian updating process (refer to equation 4.1 below) to update the conditional weights, reflecting the independence of the parameter sampling. The posterior from one application becomes the prior for the next application. As highlighted by Beven & Freer (2001b), the performance measures for a given parameter set for

the periods may be correlated, indeed it should be hoped (if not necessarily expected) that if a model performs well in one calibration period, it will continue to perform well in other periods. If this is not the case then its combined performance measure will be reduced.

**Table 4-1: Metrics used to evaluate model performance.**

Name	Equation	Notes
NSE	$\left(1 - \frac{\sigma_{\epsilon}^2}{\sigma_o^2}\right)$ <p>where <math>\sigma_{\epsilon}^2</math>, <math>\sigma_o^2</math> are the error variance and variance of the observations respectively</p>	<p>NSE is a commonly used hydrological statistical measure that measures 'goodness of fit' between a modelled and observed hydrograph (Nash and Sutcliffe, 1970)</p> <p>NSE examines the difference, every day of the record, between the simulated and observed flow value.</p>
LogNSE		<p>Similar to NSE but with the natural log of flows used to reduce the influence of large residuals near the hydrograph peak (flood flows), so better suited for water resources applications (but see (Pushpalatha et al., 2012))</p>
FDC_full	$\frac{1}{P} \sum_{p=1}^P ((Qm_p - Qo_p) / Qo_p)^2$ <p>where <math>Qm_p</math>, <math>Qo_p</math> are modelled and observed flow percentiles at flow percentiles Q1, Q5, Q10, Q30, Q50, Q70, Q90, Q95 and Q99.</p>	<p>Skill of model in reproducing the full flow duration curve based on 9 representative flow percentiles. Lower values indicate higher relative skill.</p>
FDC_low	$\frac{1}{P} \sum_{p=1}^P ((Qm_p - Qo_p) / Qo_p)^2$ <p>where <math>Qm_p</math>, <math>Qo_p</math> are modelled and observed flow percentiles at flow percentiles Q50, Q70, Q90, Q95 and Q99.</p>	<p>Skill of model in reproducing the lower section of the flow duration curve based on 5 representative flow percentiles. Lower values indicate higher relative skill.</p>
Volume_Error (PBIAS after Gupta et al., 1999)	$\frac{\sum_{i=1}^y (Qo_i - Qm_i)}{\sum_{i=1}^y Qo_i}$ <p>where <math>Qm_i</math>, <math>Qo_i</math> are modelled and observed flows on day i</p>	<p>Skill in reproducing the overall mass balance when compared to the historical flow record. Lower values indicate higher performance, with 0 indicating a perfect match, negative values indicating under estimation and positive values an over estimation (Moriassi et al., 2007)</p>
Drought_error (after Thames Water, 2010)	$\sum_{i=1}^y \left( \frac{Qm_i - Qo_i}{Qo_i} \right)^2$ <p>where <math>Qm_i</math>, <math>Qo_i</math> are modelled and observed mean summer flows for 'drought' year i.</p>	<p>Skill in reproducing mean summer flow (April to September) for the y lowest years in the record. Lower values indicate better relative skill. Informed by metric used in Thames Water Drought Plan (Thames Water, 2010)</p>

$$L(\theta_i|Y) = L_1(Y|\theta_i)L_0(\theta_i) / C \quad (4.1)$$

where  $L_0(\theta_i)$  is a prior likelihood [performance] measure for the parameter set  $\theta_i$ ;  $L_1(Y|\theta_i)$  is the likelihood [performance] measure calculated for the simulation of observed variable  $Y$  by the parameter set  $\theta_i$ ;  $L(\theta_i|Y)$  is the posterior likelihood for the parameter set  $\theta_i$  given the new observations  $Y$ ; and  $C$  is a scaling constant (to ensure that the cumulative posterior likelihood is unity)

A uniform distribution was attributed to each parameter and pre-specified threshold values were defined (used as a boundary for scaling, and below which a conditional score of 0 was used) with a shaping factor, after Freer et al. (1996), equal to 1. A feature of this use of equation 4.1 is that if a model is non-behavioural on any evaluation then the posterior likelihood (conditional score) for that model will be zero regardless of how well it performs elsewhere.

The threshold values were deliberately set conservatively with the requirement that a parameter set would have to meet the minimum threshold for all measures to produce a non-zero score for each evaluation period. The thresholds also considered the values used in similar studies presented in Chapter 2 (e.g. Cloke et al., 2010; Christerson et al., 2012; Matrosov et al., 2013; Borgomeo et al., 2014) and the model evaluation guidelines proposed by Moriasi et al. (2007). For the Nash-Sutcliffe Efficiency measure, a threshold value of 0.5 was adopted, for volume error (also known as PBIAS) a threshold of +/- 20% was used and for the Flow Duration Curve a threshold equivalent to +/-20% error at each measured flow percentile was used. Finally, for the 'Drought\_Error' metric, a threshold equivalent to an error of +/-20% at each value considered was used. Each performance score was then rescaled to a value between 0 and 1 by dividing the score's maximum possible value.

The prediction bounds from the resulting ensemble will be conditional on these assumptions (Beven & Binley, 2013) including the choice of measures, the associated thresholds and how they are combined (e.g. using Bayes' Equation here) which are all subjective decisions. Further, the measures may include an aspect 'of double counting' in that they may reinforce similar aspects of model behaviour. However, even where independent measures can be identified, the relative weights applied to these would incorporate some aspect of subjectivity - it is difficult, a priori, to understand the relative importance of different measures on the final objectives of the modelling (in this case the projected supply – demand balance for London). However, similarly to when considering multiple time-periods, the performance measures for a given parameter set for a given metric may be correlated, and that if a model performs well according to one metric, it will also perform well under other metrics. If this is not the case then its combined performance



measure will be reduced. However, the basis of this evaluation criteria and the overall approach to characterising uncertainty must be considered when interpreting and reviewing subsequent ensemble predictions.

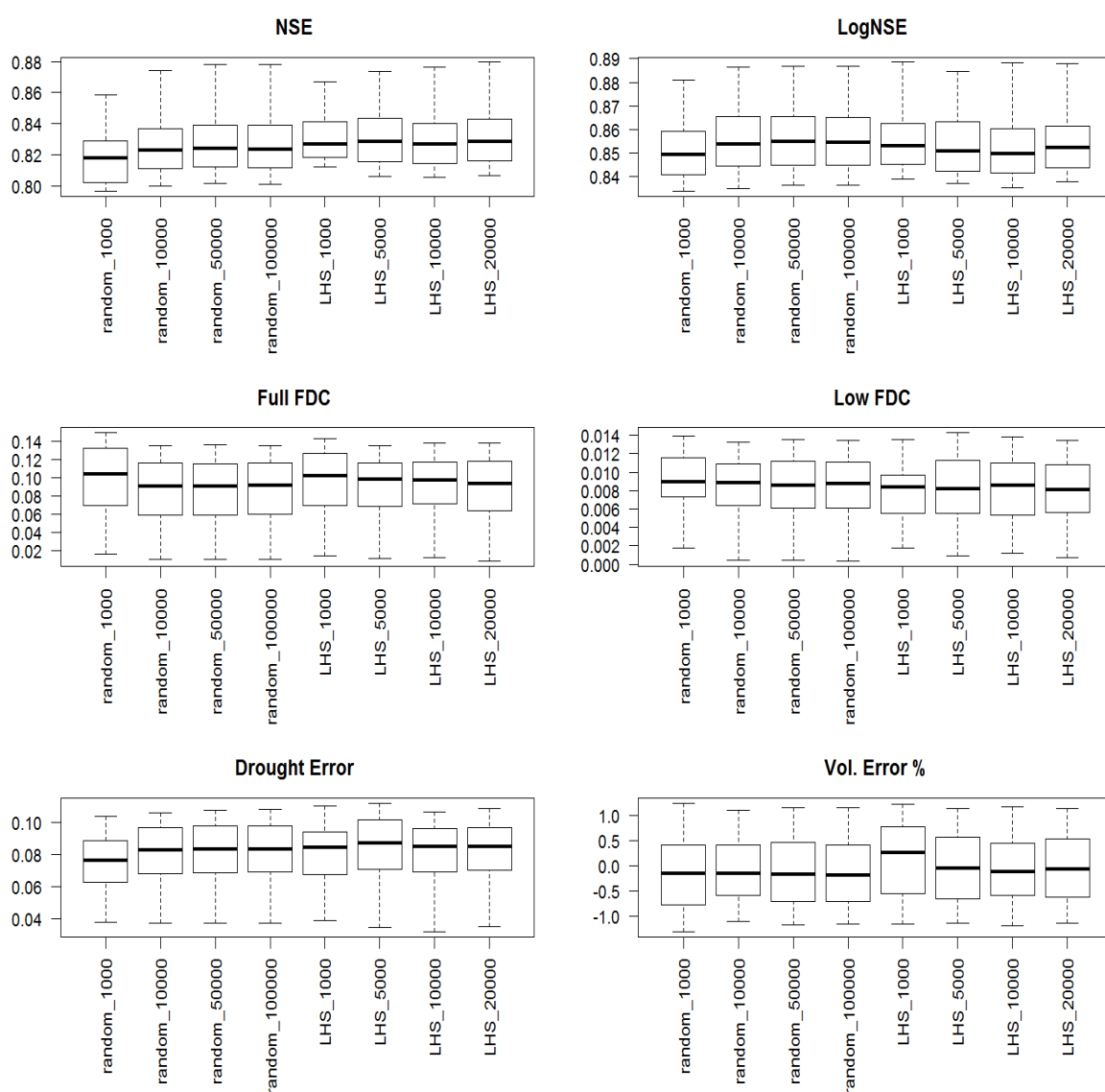
#### 4.3.1.3 Hydrological model parameterisation and sampling scheme

Similarly to the choice of performance criteria and hydrological model structures, there is an element of subjectivity in the choice of upper and lower bounds for the hydrological model parameter ranges as well as the approach to the sampling of the resulting uncertainty space. For this research wide parameter ranges, informed from user-manuals and previous studies (Moore and Bell, 2002; UKWIR, 2007; Moore et al., 2007; Wilby & Harris, 2006) were used to define the limits of the parameter space for each model structure.

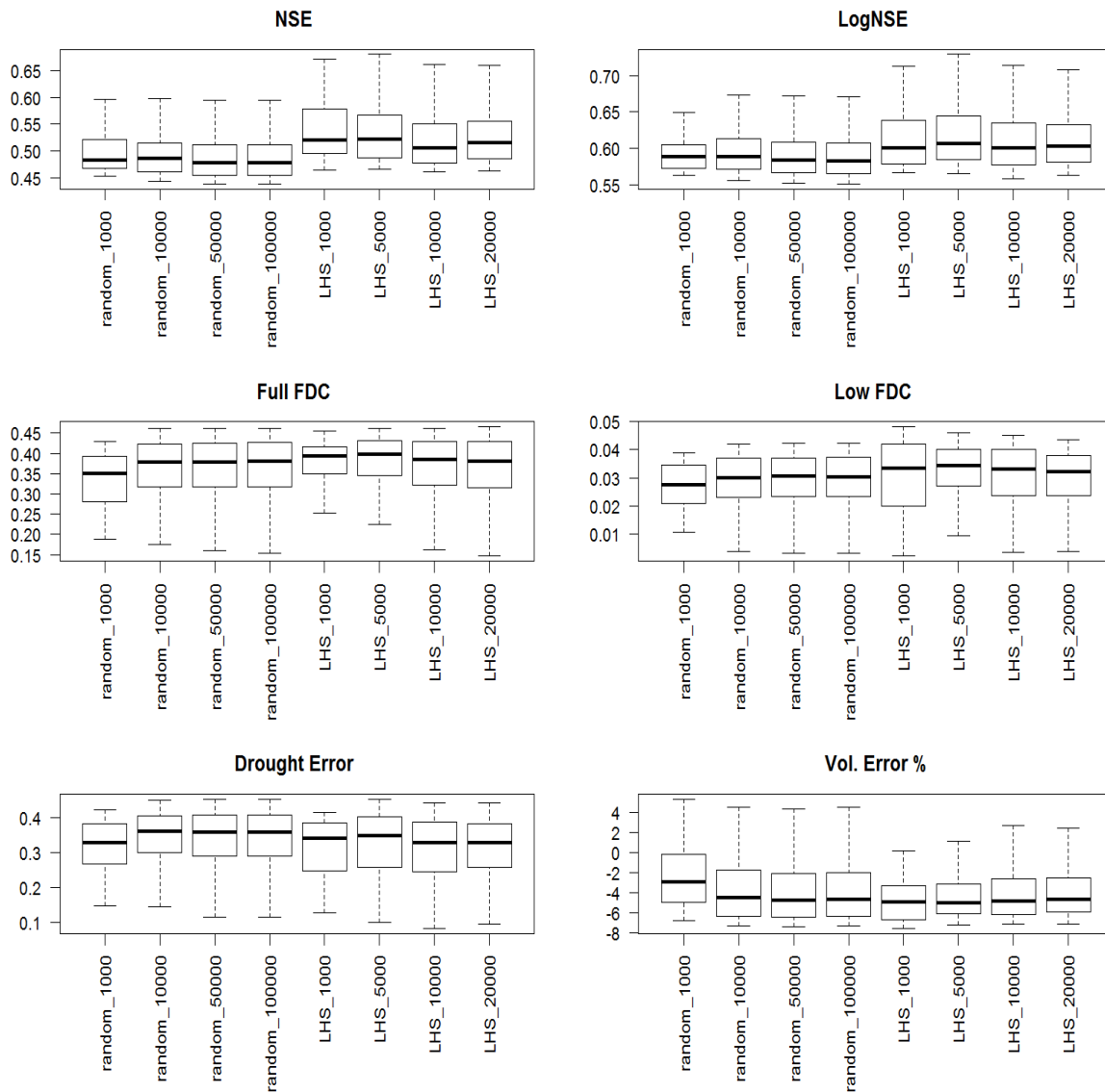
For the sampling, a random sampling strategy is commonly used in research studies adopting a GLUE approach because it has the advantage of, where a sufficiently large sample is taken, systematically exploring the uncertainty space within the upper and lower bounds provided. However, a key design consideration was ensuring an effective and practicable computational modelling approach was developed for the research. To reduce the computational burden whilst maintaining adequate representation of the full uncertainty space, a Latin Hypercube Sampling (McKay et al., 1979) approach was explored.

Latin Hypercube Sampling (LHS) seeks to divide the parameter space into regions of equal probability and consequently generate a sample of plausible collections of parameter values from a multidimensional distribution (McKay et al., 1979). LHS has been applied in a number of previous hydrological studies (e.g. Murphy et al., 2006; Arnold et al., 2009; Chaney et al., 2015; Broderick et al., 2016). The effectiveness of LHS was considered in the context of hydrological modelling by Yu et al. (2001) which reported that of a number of alternative methods tested only a LHS approach produced results similar to adopting a random sampling approach.

As detailed in Beven & Binley (2013) it is not always clear just what sort of density of sampling results from a specific sampling strategies and the approach in effect defines the prior information at the outset. In this research, the use of the Latin Hypercube Sampling was compared with random sampling ensemble for the Thames catchment draining to Kingston. This compared the distribution of performance (for the Thames catchment draining to Kingston) across the six water resources relevant measures to increasing sizes of randomly and LHS sampled ensembles over the baseline period of 1961 to 2008. Figure 4-1 and Figure 4-2 summarise the distributions for the top 5% of models in each ensemble for both the PDM-AQ and TCM model structures respectively.



**Figure 4-1: Distribution of PDM-AQ model performance (top 5% of sampled model parameter sets), as measured by the six water resources relevant performance measures (see Chapter 3), for the Thames catchment to Kingston. Performance scores (not rescaled) calculated for the baseline period 1961 – 2008 using the Oudin PET method.**



**Figure 4-2: Distribution of TCM model performance (top 5% of sampled model parameter sets), as measured by the six water resources relevant performance measures (see Chapter 3), for the Thames catchment to Kingston. Performance scores (not rescaled) calculated for the baseline period 1961 – 2008 using the Oudin PET method.**

Based on these results, the performance distributions of Latin Hypercube Sample (LHS) sizes of 5,000, 10,000 and 20,000 were considered sufficiently similar to that achieved by the 100,000 randomly sampled sets. Significant differences in the distributions are most evident when considering only 1,000 randomly sampled parameter sets and it was concluded that identifying an ensemble of models from a starting ensemble of 10,000 LHS parameter sets per catchment would be a reasonable compromise between ensuring effective sampling coverage and density and the computational resource that would be required in later parts of the research. It is noted that similarly sized ensembles using a LHS approach were adopted by Broderick et al. (2016).

#### 4.3.1.4 Hydrological model performance through evaluation periods

For this research, for a model to be retained (not rejected) as part of the final ensemble, its performance across five separate simulation periods was required to satisfy the six performance criteria set-out in Section 4.3.1.1, with the performance scores combined for each performance measure following a Bayes' updating process (see Section 4.3.1.2).

Table 4-2 presents the number of simulations that achieved a non-zero overall score for each individual simulation period for each combination of sub-catchment and model structure, considering only those simulations run using the Oudin method for estimating PET. The results in Table 4-2 show that, across all sub-catchments, far more PDM-AQ models achieved a non-zero performance score than for the TCM model structure. Both model structures achieved fewer non-zero performance scores for catchment 39011 (Tilford) and the greatest number of non-zero performing models for catchment 39016 (Theale) and 39001 (Kingston). The catchments 39016 (Theale) and 39001 (Kingston) are the larger of the catchments and both also contain significant proportions of each hydrogeological class (see Table 3-1) - and hence the greatest number of parameters significantly influencing model performance. Catchment 39021 (Enslow Mill) is the most flashy of the catchments (in that it has the greatest proportion of the catchment with an 'aquitard' hydrogeological classification) and therefore the performance score (particularly the NSE based measures) may be more sensitive to the reproduction of the timing and magnitude of flood peaks.

Table 4-3 shows the number of models that are retained after each simulation period is added in chronological order following the Bayes' updating process (see Section 4.3.1.2). These results demonstrate the impact of additional information content in assessing a model's performance can significantly reduce the number of models retained. This is particularly evident when looking at two contrasting periods, the relatively dryness between 1971 and 1980 and the relatively wetness between 2001 and 2008. The TCM model structure fails to deliver any models for catchments 39006 (Newbridge) and 39011 (Tilford) when considering all simulation periods.

**Table 4-2: Number of models (out of 10,000 LHS parameter sets) that achieved a non-zero score during each of the five separate simulation time-periods. Models driven using the Oudin (Oudin et al., 2005) method for estimating PET only are presented. Catchment locations are shown in Figure 3-1.**

<b>Catchment / Time Period</b>	<b>1961 - 1970</b>	<b>1971 – 1980</b>	<b>1981 - 1992</b>	<b>1993 – 2000</b>	<b>2001 - 2008</b>
<b>39001 (Kingston)</b>	6935 (PDM-AQ) 1011 (TCM)	5500 (PDM-AQ) 792 (TCM)	5065 (PDM-AQ) 671 (TCM)	5617 (PDM-AQ) 2426 (TCM)	5283 (PDM-AQ) 794 (TCM)
<b>39006 (Newbridge)</b>	2333 (PDM-AQ) 170 (TCM)	114 (PDM-AQ) 12 (TCM)	460 (PDM-AQ) 19 (TCM)	836 (PDM-AQ) 41 (TCM)	1027 (PDM-AQ) 21 (TCM)
<b>39011 (Tilford)</b>	1202 (PDM-AQ) 2 (TCM)	684 (PDM-AQ) 37 (TCM)	1420 (PDM-AQ) 140 (TCM)	1097 (PDM-AQ) 299 (TCM)	596 (PDM-AQ) 4 (TCM)
<b>39016 (Theale)</b>	4029 (PDM-AQ) 971 (TCM)*	5374 (PDM-AQ) 1204 (TCM)	6234 (PDM-AQ) 3668 (TCM)	5861 (PDM-AQ) 2273 (TCM)	5078 (PDM-AQ) 2437 (TCM)
<b>39021 (Enslow Mill)</b>	Not assessed**	86 (PDM-AQ) 22 (TCM)	1527 (PDM-AQ) 575 (TCM)	1924 (PDM-AQ) 2244 (TCM)	1824 (PDM-AQ) 466 (TCM)

\* record starts in 1962

\*\* record starts in 1967

**Table 4-3: Remaining number of models (out of 10,000 LHS parameter sets) achieving a non-zero performance score as the cumulative information content from each simulation period is added in chronological order – models driven using the Oudin method for estimating PET only are presented.**

<b>Catchment / Time Period</b>	<b>1961 - 1970</b>	<b>1971 – 1980</b>	<b>1981 - 1992</b>	<b>1993 – 2000</b>	<b>2001 - 2008</b>
<b>39001 (Kingston)</b>	6935 (PDM-AQ) 1011 (TCM)	4831 (PDM-AQ) 189 (TCM)	4007 (PDM-AQ) 134 (TCM)	3803 (PDM-AQ) 134 (TCM)	2421 (PDM-AQ) 18 (TCM)
<b>39006 (Newbridge)</b>	2333 (PDM-AQ) 170 (TCM)	15 (PDM-AQ) 0 (TCM)	15 (PDM-AQ) 0 (TCM)	15 (PDM-AQ) 0 (TCM)	15 (PDM-AQ) 0 (TCM)
<b>39011 (Tilford)</b>	1202 (PDM-AQ) 2 (TCM)	202 (PDM-AQ) 2 (TCM)	202 (PDM-AQ) 2 (TCM)	94 (PDM-AQ) 2(TCM)	11 (PDM-AQ) 0 (TCM)
<b>39016 (Theale)</b>	4029 (PDM-AQ) 971 (TCM)*	2875 (PDM-AQ) 126 (TCM)	2818 (PDM-AQ) 126 (TCM)	2814 (PDM-AQ) 126 (TCM)	2758 (PDM-AQ) 126 (TCM)
<b>39021 (Enslow Mill)</b>	Not assessed**	86(PDM-AQ) 22 (TCM)	67(PDM-AQ) 10 (TCM)	59 (PDM-AQ) 10 (TCM)	25 (PDM-AQ) 1 (TCM)

\* record starts in 1962

\*\* record starts in 1967

Whilst several previous studies have made use of the TCM model structure (e.g. Cloke et al., 2010; Christiersen et al., 2012), particularly in the Thames (e.g. Wilby & Harris, 2006; Manning et al., 2009; Borgomeo et al., 2014) for water resources investigations, the results presented here suggest that this model structure may not be particularly appropriate for providing predictions of river flow at all sites because no models were retained using either method for estimating PET for study catchment 39006 (Newbridge) - and only one for catchment 39011 (Tilford) and five in total for catchment 39021 (Enslow Mill).

The PDM-AQ model structure includes both fast and slow pathways for the aquifer classifications whereas the TCM model structure has a single flow pathway for each hydrogeological classification. It is speculated that the TCM model structure, with its single flow pathway, is less able to reproduce the range of flow mechanisms likely to be occurring across all the catchments compared to the PDM-AQ model structure. The PDM-AQ model structure also adopts a probability distributed root store (see Section 3.3.3) which allows for variability in moisture holding capacity rather than the uniform profile presumed by the TCM model structure. The results presented here suggest that the greater flexibility of the PDM-AQ model structure appears to enable it to be more readily calibrated or ‘conditioned’ than the TCM model structure across the Thames catchment.

However, it is acknowledged that the relative performance of the model structures will be partly dependent upon the design of the GLUE based approach to considering uncertainty, including the choice of performance metrics and model conditioning periods. Each model structure, with the four HRUs used, has 20 ‘free’ parameters and further research could include a systematic diagnostic evaluation of the model structures to better understand why certain characteristics of the hydrological signature, particularly at low flows, are particularly difficult to reproduce.

#### *4.3.2 Incorporating the use of two contrasting methods for estimating PET*

The influence of the method used to calculate PET is likely to be significant (Haxton & Young, 2012), particularly at low flows in permeable lowland catchments in the south of England (Charlton & Arnell, 2014).

Table 4-4 presents the number of retained models for each study catchment across all evaluation periods comparing their performance using both model structures and both methods for estimating PET. These results highlight that a greater number of models satisfy the performance criteria using the Oudin method for catchments 39001 (Kingston) and 39011 (Tilford), with catchments 39006 (Newbridge), 39016 (Theale) and 39021 (Enslow Mill) retaining a greater number of models using the Penman – Monteith formula.

**Table 4-4: Comparison of the number of retained models (out of 10,000 LHS parameter sets), considering all evaluation periods using the two different approaches for estimating PET.**

Catchment / PET method	Catchment				
	39001 (Kingston)	39006 (Newbridge)	39011 (Tilford)	39016 (Theale)	39021 (Enslow Mill)
<b>Oudin</b>	2421 (PDM-AQ) 18 (TCM)	15 (PDM-AQ) 0 (TCM)	11 (PDM-AQ) 0 (TCM)	2758 (PDM-AQ) 126 (TCM)	25 (PDM-AQ) 1 (TCM)
<b>Penman-Monteith</b>	1262 (PDM-AQ) 0 (TCM)	47 (PDM-AQ) 1 (TCM)	3 (PDM-AQ) 0 (TCM)	3442 (PDM-AQ) 429 (TCM)	32 (PDM-AQ) 4 (TCM)

These differences, along with the findings of Section 3.4.2.1, suggest that the method used for estimating PET may be significant (as suggested by Haxton & Young, 2012), particularly during hot, dry periods, noting that during such periods actual evapotranspiration will be less than the potential rate due to crop stress arising from limited available soil moisture. Figure 4-3 compares the model performance metrics for the retained models using both methods for estimating PET using the PDM-AQ model structure. In this figure each performance score has been rescaled to a value between 0 and 1 by dividing by the maximum possible value [i.e. a perfect model] for that performance score. An overall performance score has been calculated over the entire evaluation record for each measure by taking the geometric mean (multiplying each score across each period and taking the  $n^{\text{th}}$  root where  $n$  is equal to the number of evaluation periods (i.e.  $n$  equals 5 for study catchments 39001 (Kingston), 39006 (Newbridge), 39011 (Tilford) and 39016 (Theale) and  $n$  equals 4 for study site 39021 (Enslow Mill))).

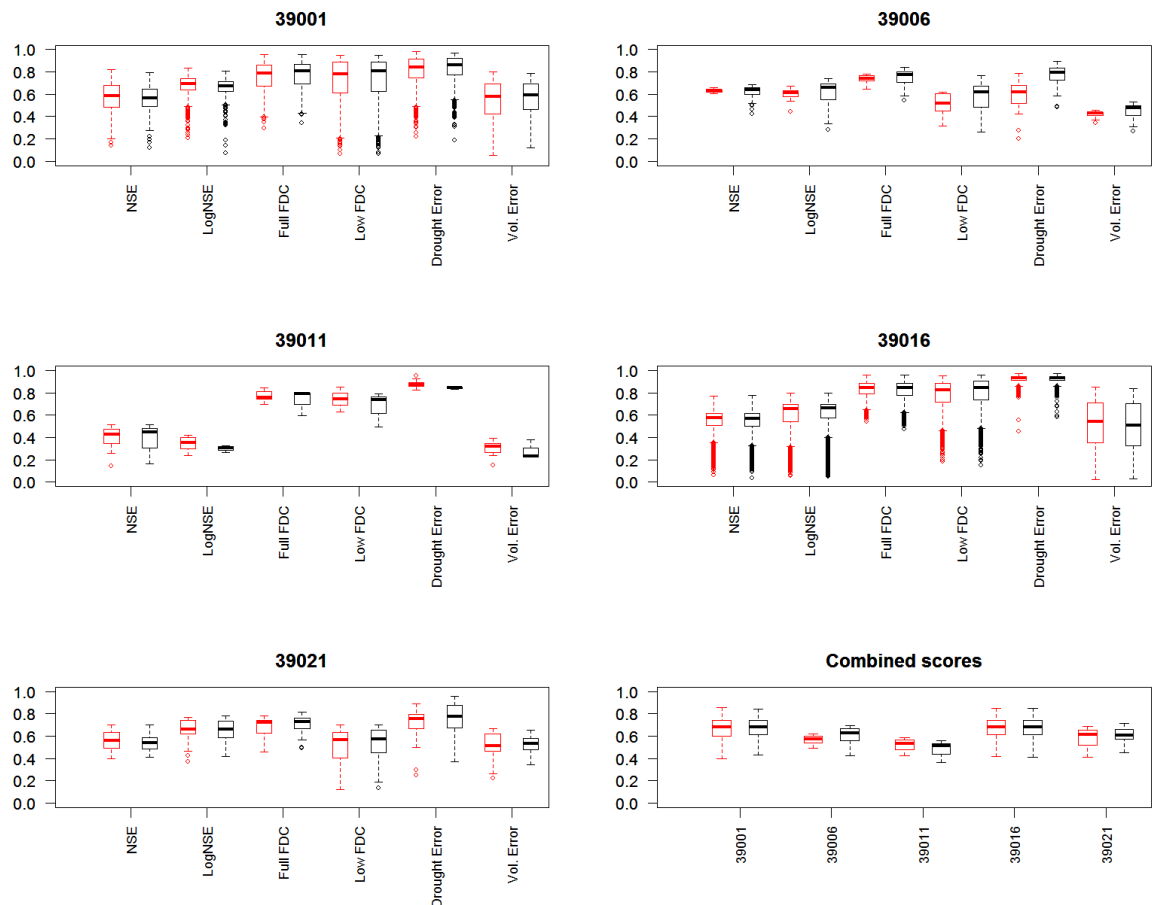
Despite the differences in the PET methods and the number of models retained, there is no evidence in this figure of a systematic difference in the scores for any of the metrics used including those such as ‘Drought error’ which might be expected to be particularly sensitive to the PET method. The potential influence of the PET method on river flow under a changed climate is investigated in Section 5.3.1.2 and in terms of water resources planning in Section 6.4.2.2).

#### 4.3.3 Exploring a gridded soil moisture accounting scheme

For the Thames catchment to Kingston (NRFA catchment 39001), the 10,000 parameter sets from the TCM HRU and PDM-AQ HRU models were also applied to a semi-distributed form of the model structures (see Section 3.3) to explore the potential influence of using the gridded climate directly as part of adopting a gridded soil moisture accounting scheme.

The simulated flows (and resulting performance measures) produced by the use of a gridded soil moisture accounting scheme showed little difference to the values produced using HRU forms of

the model structures. Consequently, it was decided that this model form would not be taken forward further as part developing the multi-site model ensemble, with the additional computational effort considered unlikely to offer significant benefits over and above nesting HRU forms of the model structures (see the following section). For areas with less homogeneity in topography, climate and hydrogeology, a semi-distributed form may offer benefits in being able to readily capture the such spatial variability in the climate and soil moisture accounting process that were not particularly evident in the modelling of the Thames catchment.



**Figure 4-3: Comparison of model performance, as measured by the six water resources relevant performance measures, for the PDM-AQ HRU catchment ensembles for each study site over all evaluation periods. Models driven using the Oudin (red) and modified Penman-Monteith (black) methods for estimating PET shown.**

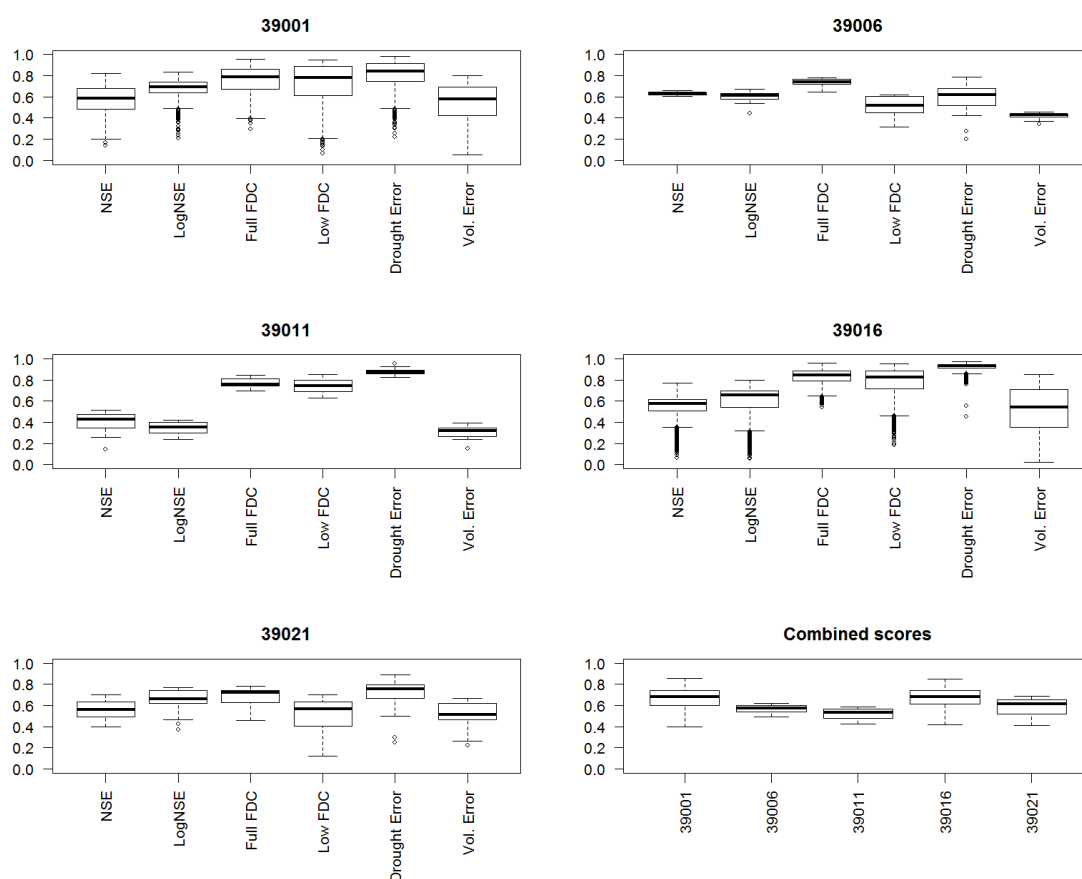
#### 4.3.4 Constructing the Multi-Site Model Ensemble (MSME)

This section describes the development of the multi-site model ensemble (MSME) used in the remainder of this research to generate spatially coherent flows across the study locations presented in Figure 3-1. Based on the above findings, the TCM model structure and the ‘gridded’ soil moisture accounting scheme form of the PDM-AQ model structure were excluded from consideration. Models driven by both methods for estimating PET were retained in developing the



final model ensemble with retained models for each study location required to achieve a non-zero score in each evaluation period (i.e. for models driven by the Oudin PET method, the retained models were those listed in the final column of Table 4-3).

Figure 4-4 shows a distribution across the different performance measures described in Section 4.3.1.1 for the retained PDM-AQ models for each study location, with the performance scores rescaled as for Figure 4-3. Figure 4-4 highlights that the thresholds used generally ensure that all performance measures contribute to a similar degree to the combined performance score. This plot also provides an indication as to which performance criteria were relatively more difficult to satisfy for a particular catchment than others. For example, the NSE and LogNSE and volume error criteria (over each time-period) are less well satisfied for catchment 39011 (Tilford) suggesting difficulties in both achieving an overall mass balance and also reproducing the exact timing of flow peaks compared to reproducing the very low summer flows during drought sequences. The narrow distributions for catchments 39006 (Newbridge) and 39011 (Enslow Mill) may be a reflection of the fewer number of parameter sets retained for these catchments.



**Figure 4-4: Comparison of model performance, as measured by the six water resources relevant performance measures, for the PDM-AQ HRU catchment ensembles for each study site over all evaluation periods. Only models driven using the Oudin method for estimating PET are shown.**

A review of the retained parameter sets for the PDM-AQ model structure across each of the locations highlighted only limited similarity (i.e. only a few parameter sets were retained in multiple locations). Consequently, a Latin Hypercube Sampling approach has been used to sample 10,000 combinations of the parameter sets for the remaining area draining to Kingston (39001) and the retained models from each of the upstream catchments. For the upstream catchments (39006 - Newbridge, 39011 - Tilford, 39016 - Theale, 39021 - Enslow Mill) the retained PDM-AQ models identified in the previous section were assigned relative weights based on their overall performance score across all evaluation periods (and performance measures) at their respective locations.

In the remainder of this thesis, this multi-site (nested) ensemble is known as the 'Multi-Site Model Ensemble (MSME)', with the model ensemble comprising only the models of the Thames basin draining to Kingston (i.e. no nested sub-models) known as the '39001 catchment ensemble'. It is recognised that this 'nesting' approach assumes independence between the retained models in the upstream catchments and that the performance at the downstream location of Kingston will be altered by the 'nesting' of models for the upstream locations. As a result, the performance at Kingston for the new nested models is re-evaluated as part of the process.

For the MSME, each sub-catchment was subjected to its own sub-catchment averaged climatology (with the remaining catchment area draining to Kingston adjusted accordingly). Flows were routed from upstream catchments to Kingston using a simple Muskingum routing process (Chow et al., 1988). With the river reaches largely baseflow dominated and flows relatively unflashy, and the focus of this research being water resources rather than capturing the timing and magnitude of flood peaks (and the hydrological modelling adopting a daily time-step), the routing parameters have been held constant (across the full flow range). A half-day travel time was specified between both Tilford (39011) and Theale (39016) to Kingston and a 1 day travel time specified between both Windrush (39006) and Enslow Mill (39021) to Kingston. The Muskingum storage weighting parameter has been fixed at a constant of 0.2 for all reaches (in natural streams the range is considered to vary between 0 and 0.3, Chow et al., 1988) with the Muskingum approach considered relatively insensitive to this parameter (Chow et al., 1988).

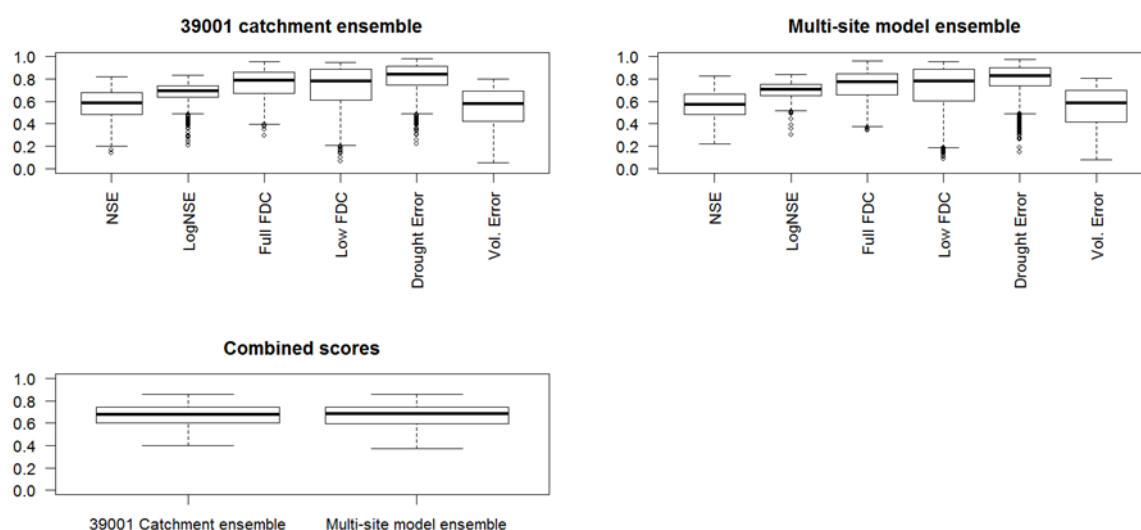
#### 4.3.4.1 Comparison of model performance from the nested hydrological models

Table 4-5 presents the number of models retained for predicting flows at Kingston from both the original model ensemble ('39001 catchment ensemble' with no nesting of upstream catchments) and the MSME.

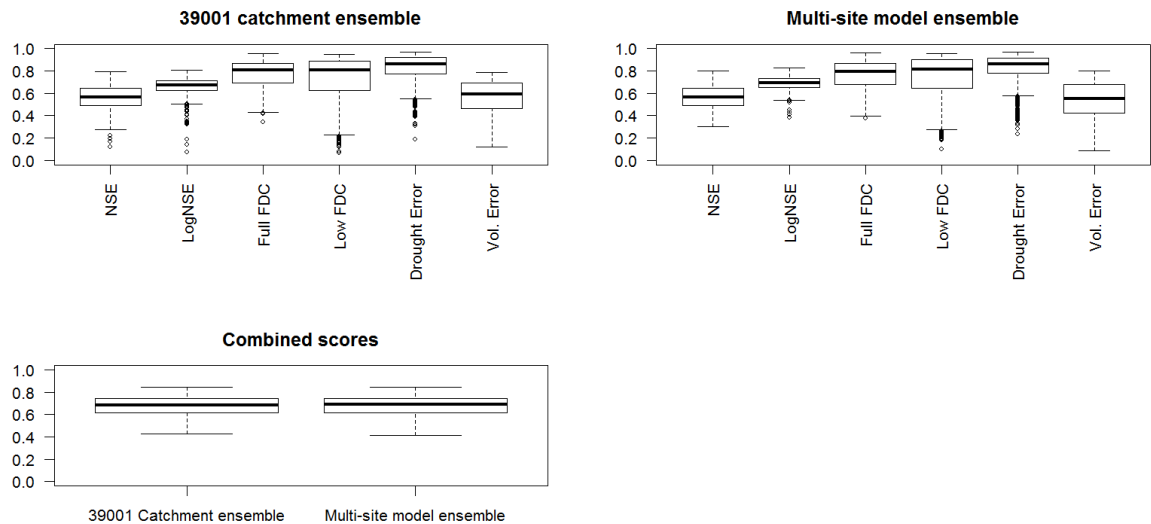
**Table 4-5: Comparison of the number of models retained (out of 10,000 LHS parameter sets) using the MSME with the '39001 catchment ensemble' for the Thames catchment to Kingston.**

PET method	Model ensemble	
	39001 catchment ensemble	MSME (nested model)
Oudin	2,421	4,215
Penman – Monteith	1,262	1,911

Figure 4-5 and Figure 4-6 show the distribution of performance for the MSME and 39001 catchment ensemble members split into the two alternative methods of estimating PET, those using the Oudin method (model ensemble known as MSME-OU, Figure 4-5) and those using the modified Penman-Monteith method (model ensemble known as MSME-PM, Figure 4-6), across the different performance measures. The plots show an equivalent performance score for each measure over the full evaluation range by multiplying the values for each measure across each period and taking the geometric mean.



**Figure 4-5: Comparison of model performance, as measured by the six water resources relevant performance measures, for the 39001 catchment ensemble and the MSME-OU (models driven using the Oudin method for estimating PET) over all evaluation periods.**



**Figure 4-6: Comparison of model performance, as measured by the six performance measures, for the original 39001 catchment ensemble and the MSME-PE (using models driven using the modified Penman-Monteith method for estimating PET) over all evaluation periods.**

These results demonstrate an increase in the number of retained models realised, along with very modest increases in performance scores, from using the MSME compared to using the 39001 catchment ensemble. The results also demonstrate that the MSME is capable of delivering models for predicting flows at Kingston as well as, by design, predictions for the contributing upstream study catchments.

#### 4.4 Assessing the predictive skill of the MSME for water resources planning

The MSME's skill in reproducing a range of hydrological characteristics of interest to the water resource planner is examined in this section along with the sensitivity of predictions to hydrological model parameter uncertainty and the two methods used to estimate PET. The hydrological characteristics considered in this evaluation of ensemble skill are the reproduction of:

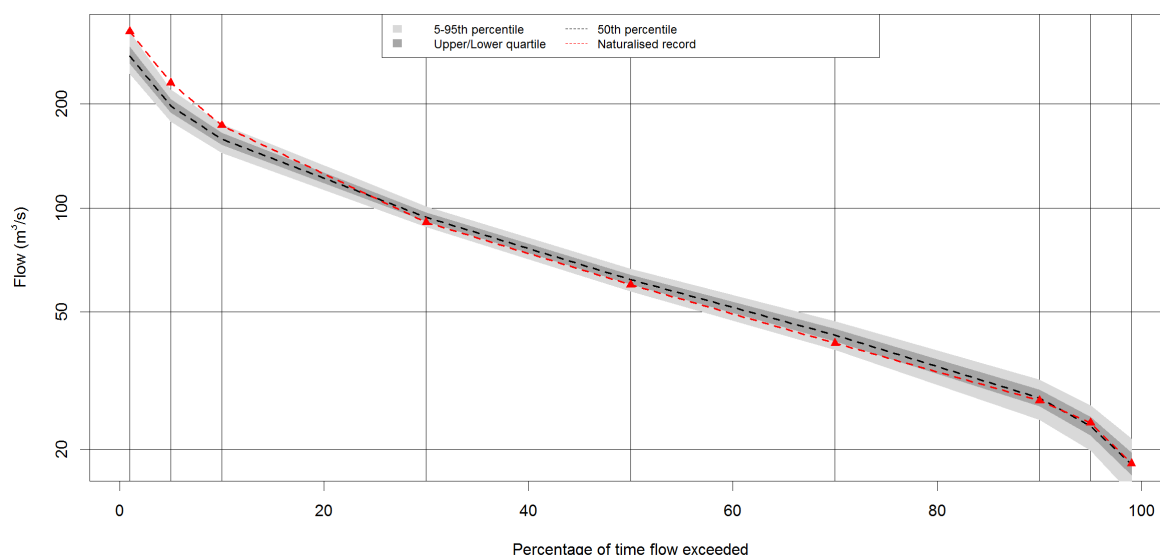
- Flow duration curves;
- Low flow / drought sequences;
- Trigger levels for the operation of strategic schemes in London, Guildford and Kennet Valley water resource zones;
- Winter flow volumes in relation to reservoir refill potential.

#### 4.4.1 River Thames at Kingston (NRFA catchment 39001)

##### 4.4.1.1 Reproduction of the Flow Duration Curve

Figure 4-7 compares the MSME's simulation of nine selected flow percentiles across the full flow regime (presented as a Flow Duration Curve (FDC)) for Kingston (NRFA catchment 39001), over the 1961 to 2008 period, to the same flow percentiles calculated using the historical flow record.

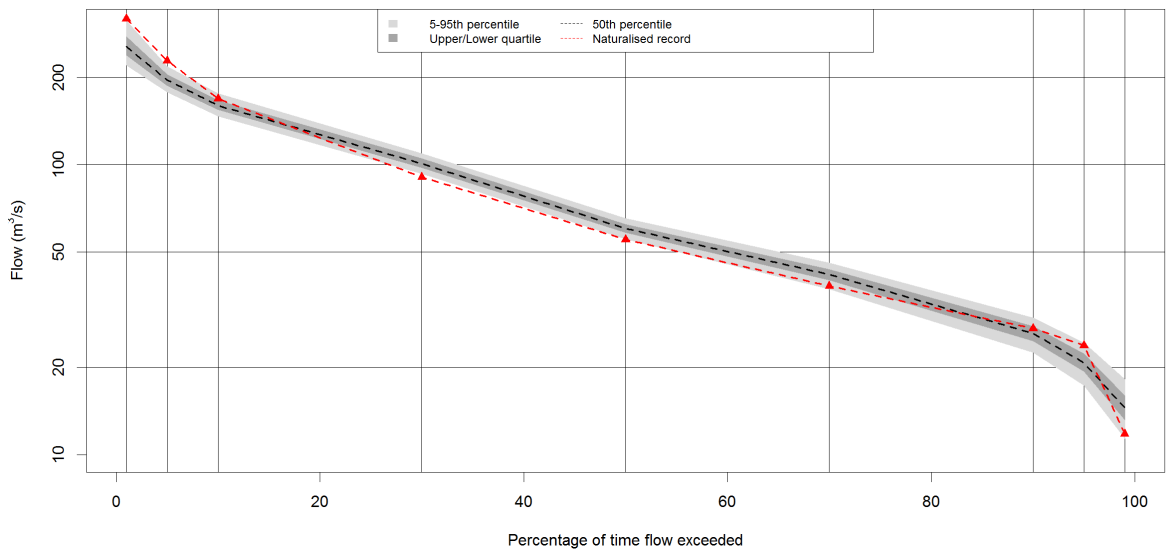
The MSME estimates (50<sup>th</sup> percentile) for the selected flow percentiles are similar to the historical record for the medium to low (Q30 to Q99) flow percentiles with large differences only apparent at the higher flows (Q1 to Q10). For Q95, often used as an indicator of low flows and widely used to characterise hydrological regimes for water management purposes (Charlton & Arnell, 2014), the predicted value varies from 19.9m<sup>3</sup>/s (5<sup>th</sup> percentile) to 26.8m<sup>3</sup>/s (95<sup>th</sup> percentile) with a central estimate from the MSME of 23.4m<sup>3</sup>/s compared to 23.9 m<sup>3</sup>/s from the historical record.



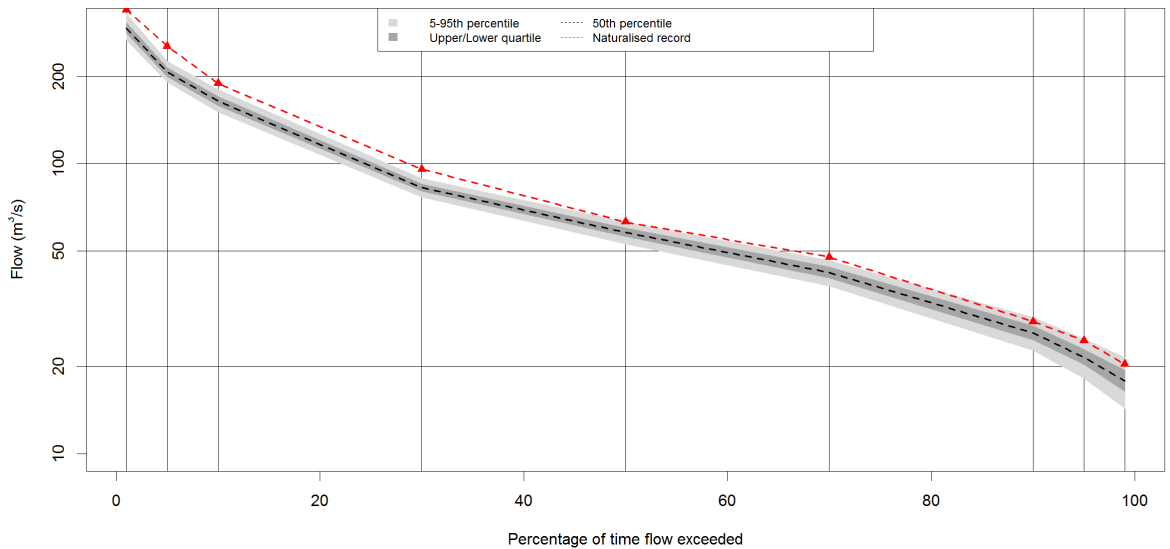
**Figure 4-7: Simulated FDC (1961 to 2008) at Kingston shown as grey shading (5<sup>th</sup> to 95<sup>th</sup> percentile estimates) and black dashed-line (50<sup>th</sup> percentile estimate). Historical FDC (1961 to 2008) is shown in red. Note that the y-axis is the natural logarithm of flows to emphasise lower flows.**

In Section 3.4, the relative dryness of the 1971 to 1980 period and, in contrast, the relative wetness of the 2001 to 2008 period in the historical flow record was highlighted. The FDCs for these two time-periods are presented in Figure 4-8 and Figure 4-9 respectively. For the 1971 to 1980 period the results demonstrate reasonable skill in prediction at the low flow percentiles (Q90, Q95 and Q99) with lower skill from Q30 to Q70 and particularly at the very high flows (Q1 and Q5). For the 2001 to 2008 period, the MSME exhibits a consistent bias in predicting lower flows than present in the historical record throughout the flow range. Examining the performance of the MSME relative to the historical record for multiple time-periods enables potential issues to

be identified that are not readily apparent when only considering the full period of simulation as presented in Figure 4-7. These contrasting time periods are examined further in Section 4.4.1.3.



**Figure 4-8: FDC (1971 to 1980) at Kingston. Simulated FDC shown as grey shading (5<sup>th</sup> to 95<sup>th</sup> percentile estimates) and black dashed-line (50<sup>th</sup> percentile estimate). Historical FDC (1971 to 1980) is shown in red.**



**Figure 4-9: Simulated FDC (2001 to 2008) at Kingston shown as grey shading (5<sup>th</sup> to 95<sup>th</sup> percentile estimates) and black dashed-line (50<sup>th</sup> percentile estimate). Historical FDC (2001 to 2008) is shown in red.**

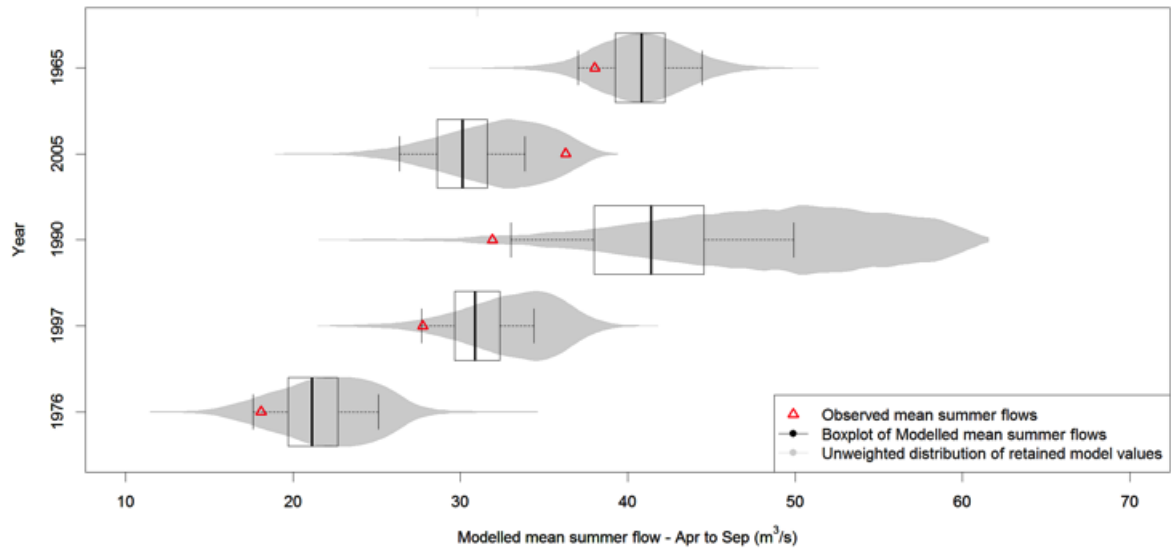
To review the performance of the MSME beyond the reproduction of selected flow percentiles, Sections 4.4.1.2 and 4.4.1.3 examine specific low-flow/drought periods and Section 4.4.1.4 considers the specific flow thresholds at which strategic schemes are activated for London. Section 4.4.1.5 considers the reproduction of winter flows with a view to the potential for additional winter storage (see Darch et al., 2011). Section 4.4.2 investigates the reproduction of hydrological characteristics at other locations in the Thames basin.

#### 4.4.1.2 Reproduction of mean summer flows in 'drought' years

Thames Water (2010) use the 'Frequency of Occurrence' (the number of years at which a given severity of drought occurs within a record) as a way of describing, in statistical terms, the severity of a particular drought. This is used as an important guide to the conformance between planned Levels of Service (see Section 6.3.4.4 for more details) and the frequency at which drought measures would be introduced.

For the London water resource zone, Thames Water use the mean summer flow (April to September) as a metric for the 'frequency of occurrence' and identifying (and ranking) the driest years in the historical record. Figure 4-10 presents the distribution of estimates from the MSME, at Kingston, for the driest summer in each of the five evaluation periods, which also correspond to the five driest sequences in the full evaluation period between 1961 and 2008 - noting that 1990 was followed by another summer with relatively low flow values, and that 1997 was preceded with a relatively dry summer, both of which are not shown.

The results show that whilst the ranking of dry years is well reproduced, the MSME demonstrates a tendency to over-estimate the mean summer flows apart for the most recent summer shown (2005). The mean summer flow during 1976 predicted by the MSME varies from 17.6 m<sup>3</sup>/s (5<sup>th</sup> percentile) to 25.1 m<sup>3</sup>/s (95<sup>th</sup> percentile) with a central estimate of 21.1 m<sup>3</sup>/s compared to 18.1 m<sup>3</sup>/s from the historical record. For 2005, the MSME varies from 26.4 m<sup>3</sup>/s (5<sup>th</sup> percentile) to 33.9 m<sup>3</sup>/s (95<sup>th</sup> percentile) with a central estimate of 30.1 m<sup>3</sup>/s compared to 36.3 m<sup>3</sup>/s from the historical record. These specific sequences are examined in more detail in the following section.



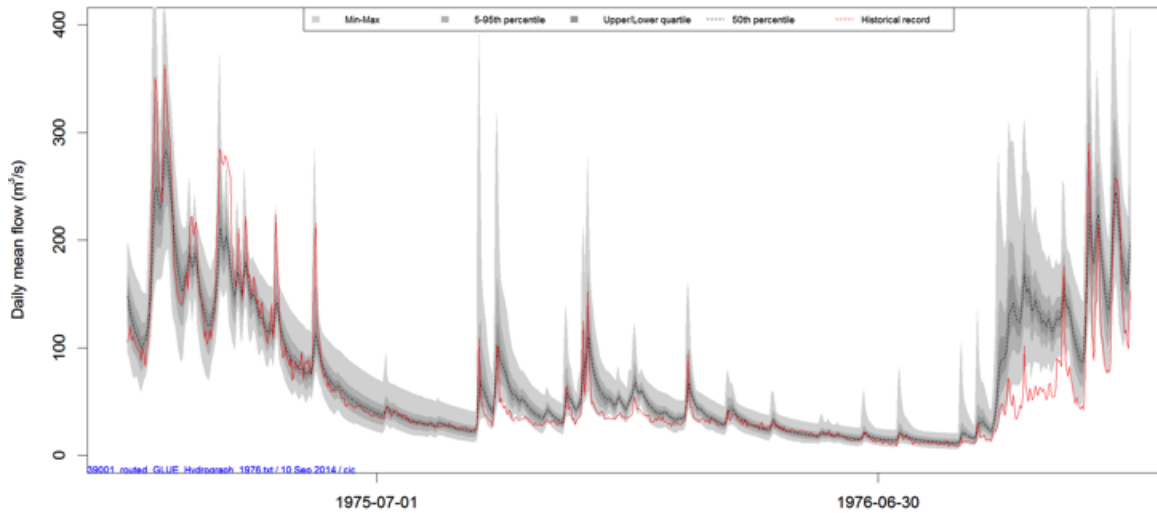
**Figure 4-10: Simulated MSF from each evaluation period at Kingston. Grey shading presents the ‘unconditioned (i.e. equally weighted)’ distribution and the black box-plots the weighted distribution with ensemble predictions conditioned by all evaluation periods (1961 to 2008 – see Section 3.4). Values from the historical record are shown as red triangles. The years shown on y-axis are the five driest summers in the historical record, ranked in terms of MSF.**

#### 4.4.1.3 Reproduction of flow hydrographs during low-flow / drought years

The previous section identified that the MSME exhibited a tendency to over-estimate flows during the summer of 1976. Consequently, Figure 4-11 presents the historical and MSME predicted flow hydrographs at Kingston during 1975 and 1976.

Figure 4-11 demonstrates that whilst the MSME is able to reproduce flows reasonably well during the summer of 1976, including the timing of the relatively small peak flows from short duration rainfall events when soil moisture deficits are high, it exhibits inconsistencies with the historical record in reproducing the wetting-up period during the autumn of 1976. It is noted in Crooks & Kay’s (2015) modelling of the Thames that their CLASSIC (a gridded derivative of the PDM model) model appeared to exhibit the same discrepancy during this period, with Crooks & Kay (2015) commenting that there is no evidence that rainfall on dry soils resulted in runoff before replenishment of the soil moisture deficit.



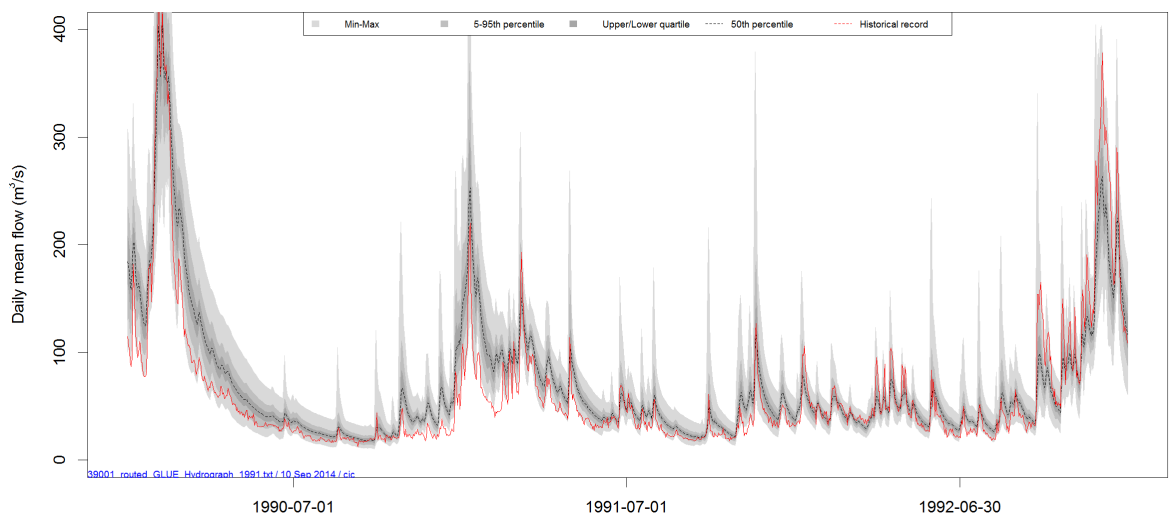


**Figure 4-11: Simulated hydrograph during 1975 and 1976 at Kingston shown in grey shading (5<sup>th</sup> to 95<sup>th</sup> percentile estimates) and black dashed line (50<sup>th</sup> percentile estimate). Values from the historical record are shown in red.**

The MSME also has a tendency to sustain flows after each of the modest summer rainfall events for longer than is evident in the historical record. Over the summer period, the hydrograph derived from the MSME used in this research confirms the ensemble's overall tendency to overestimate summer flows during 1976, as previously suggested by Figure 4-10.

To examine whether similar characteristics are exhibited during other low flow sequences, Figure 4-12 examines the historical and predicted hydrographs at Kingston during 1990 to 1992. Figure 4-10 suggests that this is the sequence that the MSME most over-estimates compared to the historical record. The mean summer flow during this period does not fall as low as 1976 but the period of relatively low flows extended over two summers as shown in Figure 4-12. The plots demonstrate that whilst the MSME is able to reproduce the wetting-up periods better than for the 1975 to 1976 period (in the autumn 1992 in particular), similar inconsistencies with the historical record to those present in 1976 are evident.

Based on the two low flow sequences, it appears that the MSME ‘releases’ more water to the river during, and just after, periods of high soil moisture deficits than is suggested by the historical flow record, when it is speculated (e.g. Oni et al., 2016) that more water may be taken up replenishing soil moisture and low aquifer levels than is predicted by the models. It is worth considering, however, that these periods are also likely to be subject to significant management and regulation of the river to maintain flows for navigation and abstraction by other users previously ‘starved’ of water. Therefore, there is likely to be relatively greater uncertainty associated with the historical flow record during such low flow periods when artificial influences will also have a greater (proportional) impact on river flows.



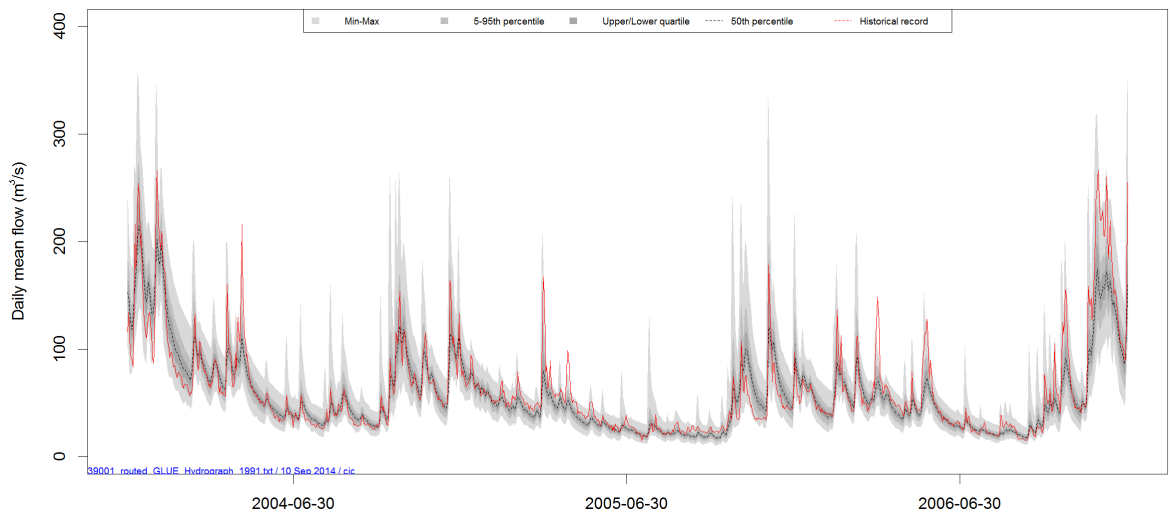
**Figure 4-12: Simulated hydrograph from 1990 to 1992 at Kingston shown in grey shading and black dashed line. Equivalent values from the historical record shown in red.**

The flow duration curves presented previously (see Figure 4-9) suggested that, in contrast to the other low flow periods, the MSME consistently under-predicts flows during the 2001 to 2008 period. Figure 4-13 examines the reproduction of the three year (2004 to 2006) sequence within this period that contains the lowest mean summer flow (2005). In contrast to the other sequences, the MSME appears to better reproduce this sequence although during the recovery in the autumn of 2006, the ensemble under-estimates flows in contrast to the sequences in 1976 and 1991/92.

Whilst a detailed examination of these historical droughts is not presented here (see Marsh et al., 2014 and Marsh et al., 2007 for more details) it is worth reflecting on some differences in these events. Marsh et al. (2014) reported that, taken together, 2005 and 2006, constitute the warmest two-year sequence in the 337-year Central England Temperature series and that the summer periods included several notably arid episodes (e.g. June 2006). Of particular interest to water resources was the overall rainfall deficiency in the winter and spring – when modest evaporation losses typically allowed the bulk of reservoir replenishment and aquifer recharge to take place. Rainfall deficiencies began to moderate through an unsettled late summer and the drought's intensity weakened substantially through the autumn of 2006 to help maintain runoff rates above drought minima.

In contrast, the hot, dry summer of 1976 was a shorter, more intense event. This is reflected in Figure 3-8 which shows the changing relative intensities of the two events when considering 12 month and 24 month accumulation time-windows in the calculation of the Standardised Precipitation and Evapotranspiration Index (SPEI). The 1976 drought was ended relatively dramatically (Marsh et al., 2014; Parry et al., 2016a) by a prolonged sequence of vigorous low pressure systems which brought heavy rainfall extending over several months in the autumn. It is speculated by the author that the physical processes in play at times of prolonged drought might be quite different to processes during other times and that the models used here were not able to fully reflect these changing dynamics.

These processes might include river reaches experiencing significant seepage to the underlying aquifer along with deep fissuring and cracking, especially of clay soils, providing new, faster, enhanced penetration of rainfall into the subsurface storage after a prolonged dry and hot period. Furthermore, Parry et al. (2016a) highlighted that abstractions from surface and groundwater sources during drought development may artificially extend the duration of the drought termination phase. These factors highlight the need for further research to understand which (or develop/improve) models might be the most suitable for the modelling of droughts for water resources planning. This challenge is exacerbated by the lack of significant drought events in the historical record against which to evaluate alternative hydrological model structures.



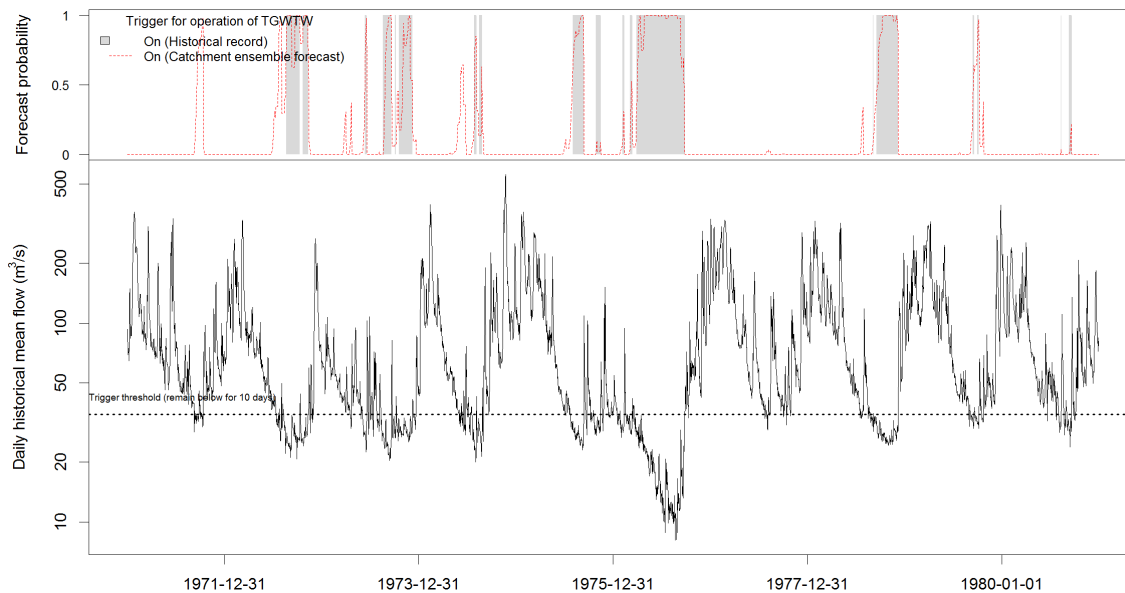
**Figure 4-13: Simulated hydrograph from 2004 to 2006 at Kingston shown in grey shading and black dashed line. Equivalent values from the historical record shown in red.**

#### 4.4.1.4 Drought Plan trigger thresholds and operation of strategic schemes

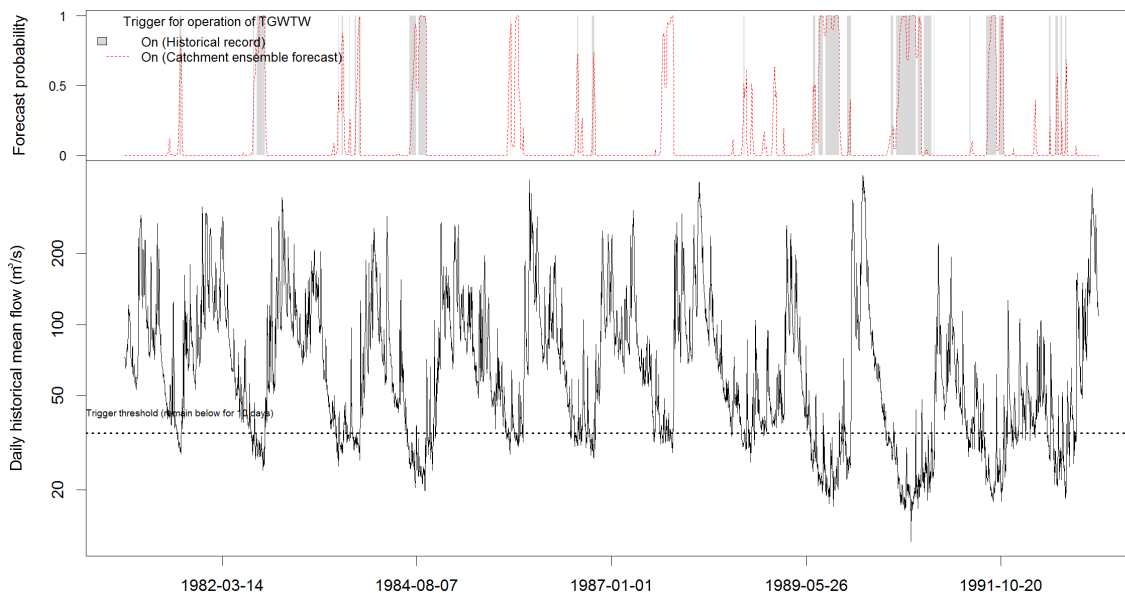
Thames Water's Drought Plan makes reference to a mean daily flow threshold in the Thames, at Kingston, of 3,000 MI/d (34.7 m<sup>3</sup>/s) for 10 or more days (Thames Water, 2010), to initiate the use of flow support measures such as the Thames Gateway desalination plant. Figure 4-14, Figure 4-15 and Figure 4-16 identify such sequences in the historical record for the three evaluation periods considered in the previous section (1971 to 1980, 1981 to 1992 and 2001 to 2008) respectively to demonstrate the performance of the MSME in its 'prediction' of such sequences.

Table 4-6 presents three measures of 'forecasting' skill to assess the performance of the MSME. It is recognised that caution should be used when comparing different evaluation periods due to the frequency of occurrence and length of each evaluation period. As a result, the measures of forecasting skill attempt to take into account their different relative frequencies of occurrence.

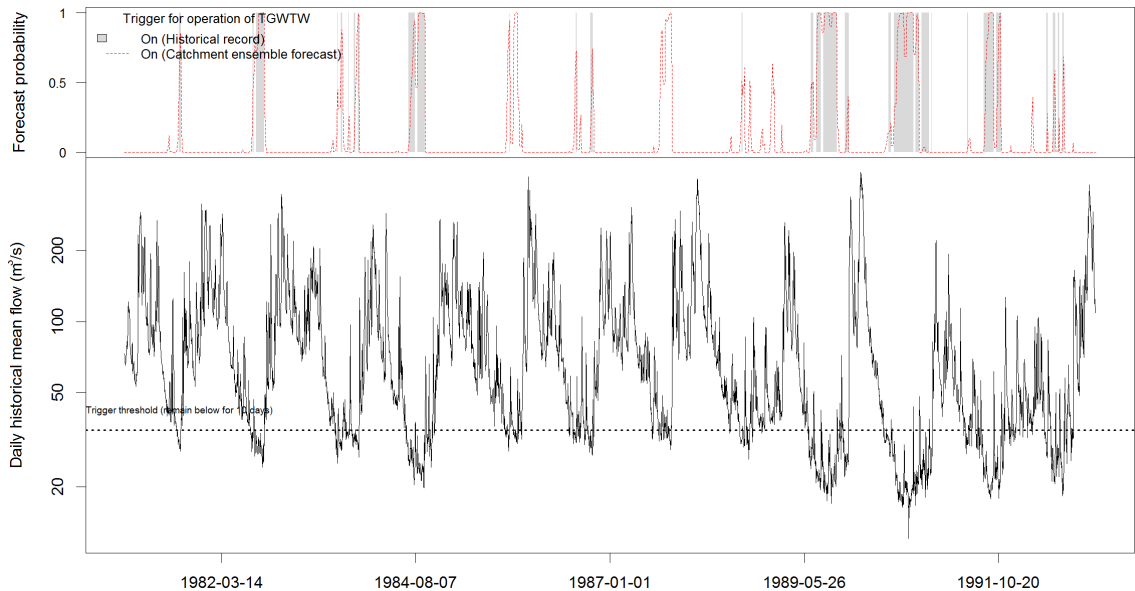
Table 4-7 summarises the number of days during each of the three evaluation periods (1971 to 1980, 1981 to 1992 and 2001 to 2008) when the trigger condition is satisfied, along with the measures of 'forecasting' skill for the MSME.



**Figure 4-14: Reproduction of trigger threshold activations at Kingston during Jan 1971 to Dec 1980 shown in red in top plot. Grey shading indicates periods in historical record where trigger threshold is crossed, with the historical flow record itself shown in black on bottom plot.**



**Figure 4-15: Reproduction of trigger threshold activations at Kingston during Jan 1981 to Dec 1992 shown in red in top plot. Grey shading indicates periods in historical record where trigger threshold is crossed, with the historical flow record itself shown in black on bottom plot.**



**Figure 4-16: Reproduction of trigger threshold activations at Kingston during Jan 1981 to Dec 1992 shown in red in top plot. Grey shading indicates periods in historical record where trigger threshold is crossed, with the historical flow record itself shown in black on bottom plot.**

The results highlight the differences in accuracy during the different time-periods. During the 1971 to 1980 and 1981 to 1992 periods the overall frequency of trigger conditions being met is close to the number in the historical record. During the 2001 to 2008 period the MSME forecasts the days when trigger conditions are met relatively well (RFAT > 0.9) but at the expense of a greater tendency to report a false positive during this period and an over estimation of the total number of days the trigger condition is satisfied compared to the other two periods. This is consistent with the MSME's tendency to underestimate flows during the 2001 to 2008 period shown in Figure 4-9. Looking at the hydrographs it is apparent that the inaccuracies are related to those periods where the historical flow is very close to the trigger threshold rather than the more extreme (both in magnitude and duration) sequences in the historical record.

**Table 4-6: Forecasting skill measures used to assess the performance of the MSME in identifying periods when trigger conditions are satisfied.**

Name	Equation	Description
Relative Forecasting Accuracy of Trigger (RFAT)	$\frac{\sum Dm}{\sum Do}$ <p>where Do is the total number of days in the historical record where the trigger condition is satisfied and Dm is the sum of the forecast probabilities on these days.</p>	<p>This assesses the degree to which the MSME correctly predicts days in the historical record that satisfy the trigger condition - a perfect score of 1 would reflect that all ensemble members correctly predict such days and a score of 0 would reflect that no members correctly predict such days.</p>
Relative False Forecasting Accuracy of Trigger (RFFAT)	$1 - \frac{\sum Dfm}{\sum Do}$ <p>where Do is the total number of days in the historical record where the trigger condition is satisfied and Dfm is the sum of forecast probabilities on days when the trigger condition is not satisfied in the historical record.</p>	<p>This assesses the degree to which the MSME predicts ‘false’ positives (trigger condition is predicted but not present in the historic record).</p> <p>The score is relative to the number of trigger days in the historical record to reflect the relative rarity of the event.</p> <p>A perfect score of 1 would reflect that all ensemble members do not predict any false positives.</p>
Brier Skill Score (BSS)	$1 - \frac{Bs}{Bs_{ref}}$ <p>where BS is the Brier score (Brier, 1950) for the MSME forecast and <math>BS_{ref}</math> is the Brier score using the frequency of occurrence in the historical record as a fixed probability estimate.</p>	<p>This form of the Brier Skill Score reflects the improvement in the accuracy of prediction compared to using the frequency of the event in the historical record as a fixed probability estimate.</p> <p>A value above 0 indicates a better predictive performance than the reference, up to a maximum score of 1 (perfect predictive skill).</p>

**Table 4-7: MSME skill in predicting periods when the trigger condition is satisfied in the historical record.**

	Evaluation Period		
	1971 – 1980 (10 years)	1981 – 1992 (12 years)	2001 – 2008 (8 years)
<b>No. of days trigger on in historical record</b>	556	527	313
RFAT	0.78	0.67	0.91
RFFAT	0.80	0.67	0.52
BSS	0.69	0.48	0.59

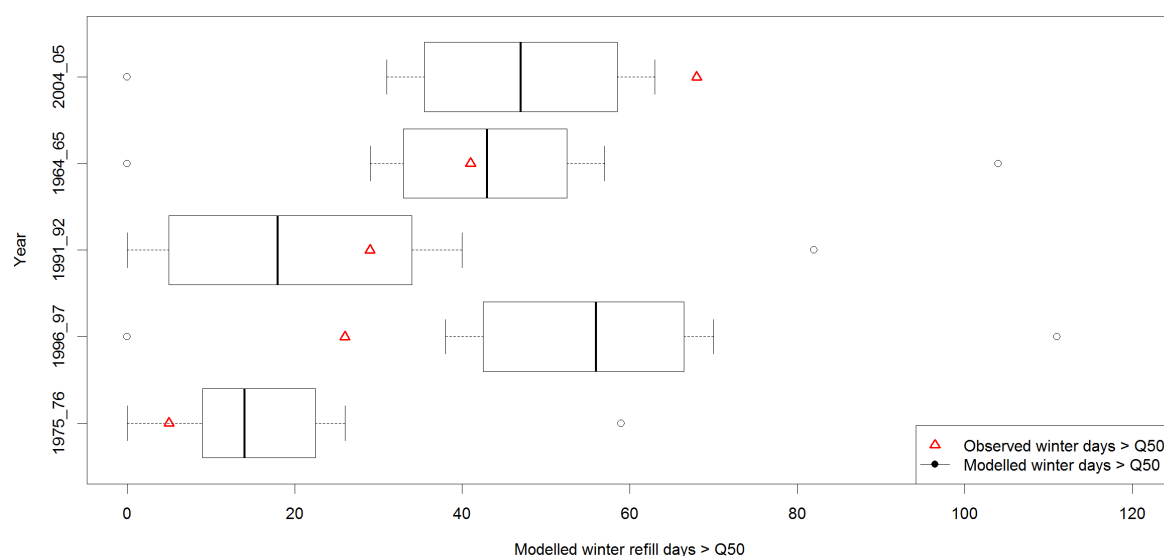
#### 4.4.1.5 Winter refill flows

Darch et al. (2011) suggested that the average number of days per month between December and May with flows at, or above, Q50, is a reasonable indicator of assessing the refill potential of a new reservoir in the Upper Thames. Figure 4-17 presents the predictions from the multi-site model ensemble in terms of the number of days above Q50 (for each model this is relative to its own prediction of Q50) and demonstrates that the MSME broadly reproduces the relative ranking of years reasonably well but highlights particular inconsistencies compared to the historical record in the winter of 1996/1997 and 2004/2005 and, to a lesser extent, 1975/1976.

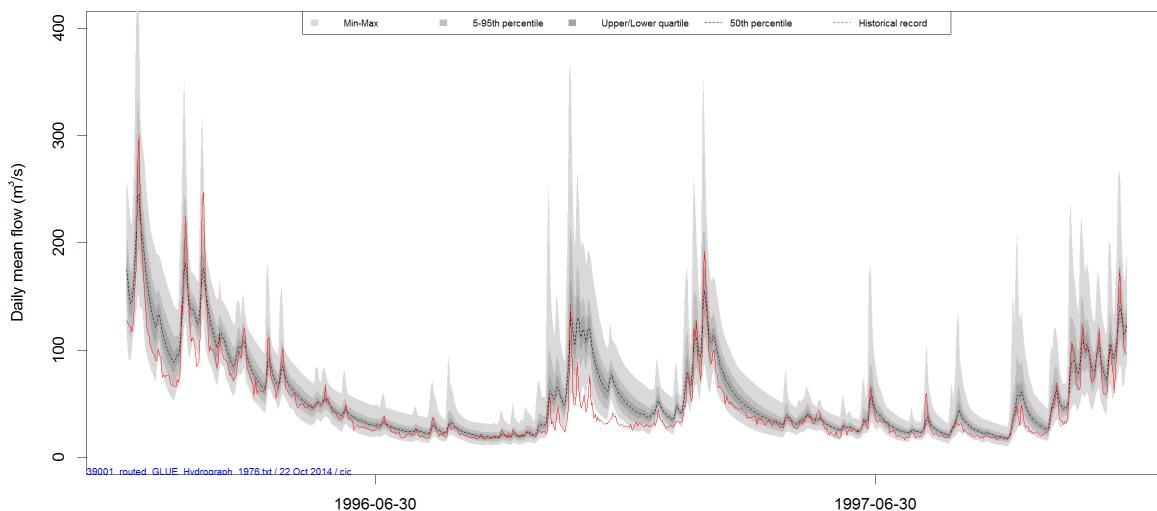
The hydrograph for 1996/1997 is presented in Figure 4-18 and highlights this issue in more detail during the winter of 1996. The summer of 1996 (and 1995 to a lesser extent) was particularly dry so this issue may be related to the ‘wetting-up’ problem identified when looking at low summer flows in Section 4.5.2. During such periods the MSME may tend to ‘release’ more water to the river during, and just after, periods of high soil moisture deficits than is realised in reality where more water may be taken up replenishing soil moisture and low aquifer levels.



It is also observed that the range of predictions includes some models which predict no days exceeding Q50 during these winters. The performance values used in this research have been derived using a range of measures which include criteria that specifically target lower flows such as the drought-error metric. Therefore using these performance values for higher flow prediction may not be ideal in promoting those models most suitable for considering winter flows. It would be interesting to explore whether this relative accuracy of the ensemble members during different periods could be reflected in their use in prediction to see if improved reproduction of the historical flow record across a wider range of conditions could be achieved. However, such an approach would present challenges, both in identifying when to switch between predictive modes and in the subsequent use of the resulting flow series in water resources system modelling such as that presented in Chapter 6 (where temporal coherence in each simulated river flow series is typically necessary).



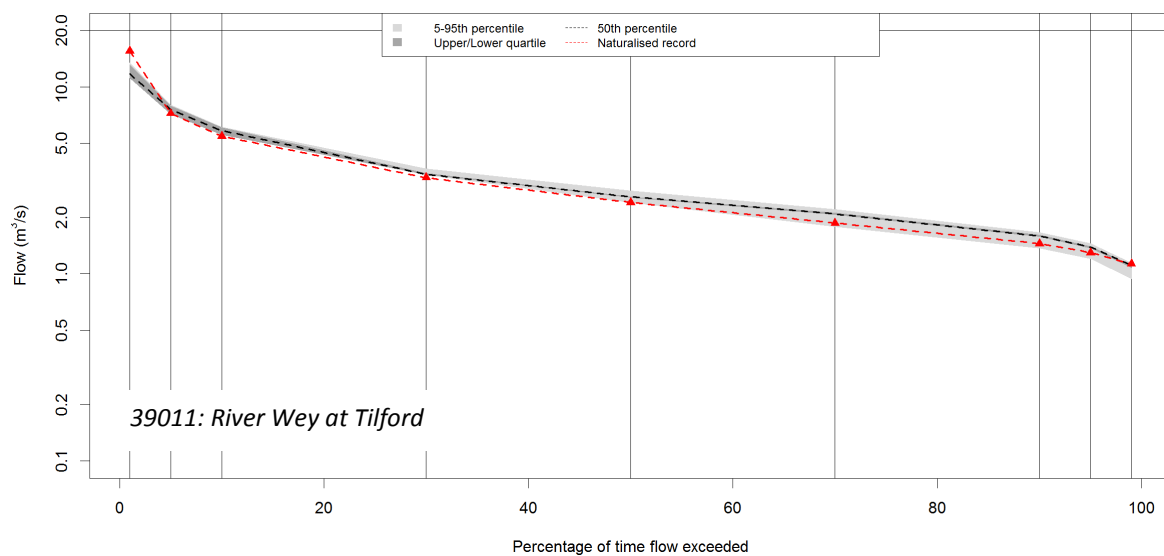
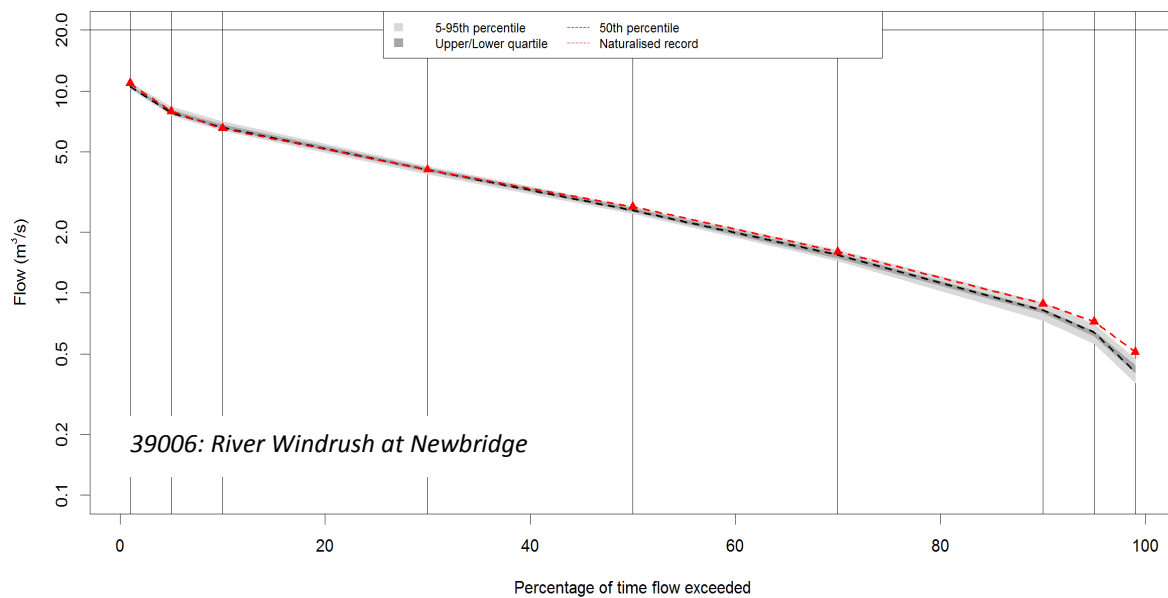
**Figure 4-17: Reproduction of lowest number of days in winter above Q50 from each evaluation period at Kingston. Note that black circles are the full extent of the modelled values. Equivalent values from the historical record shown as red triangles.**



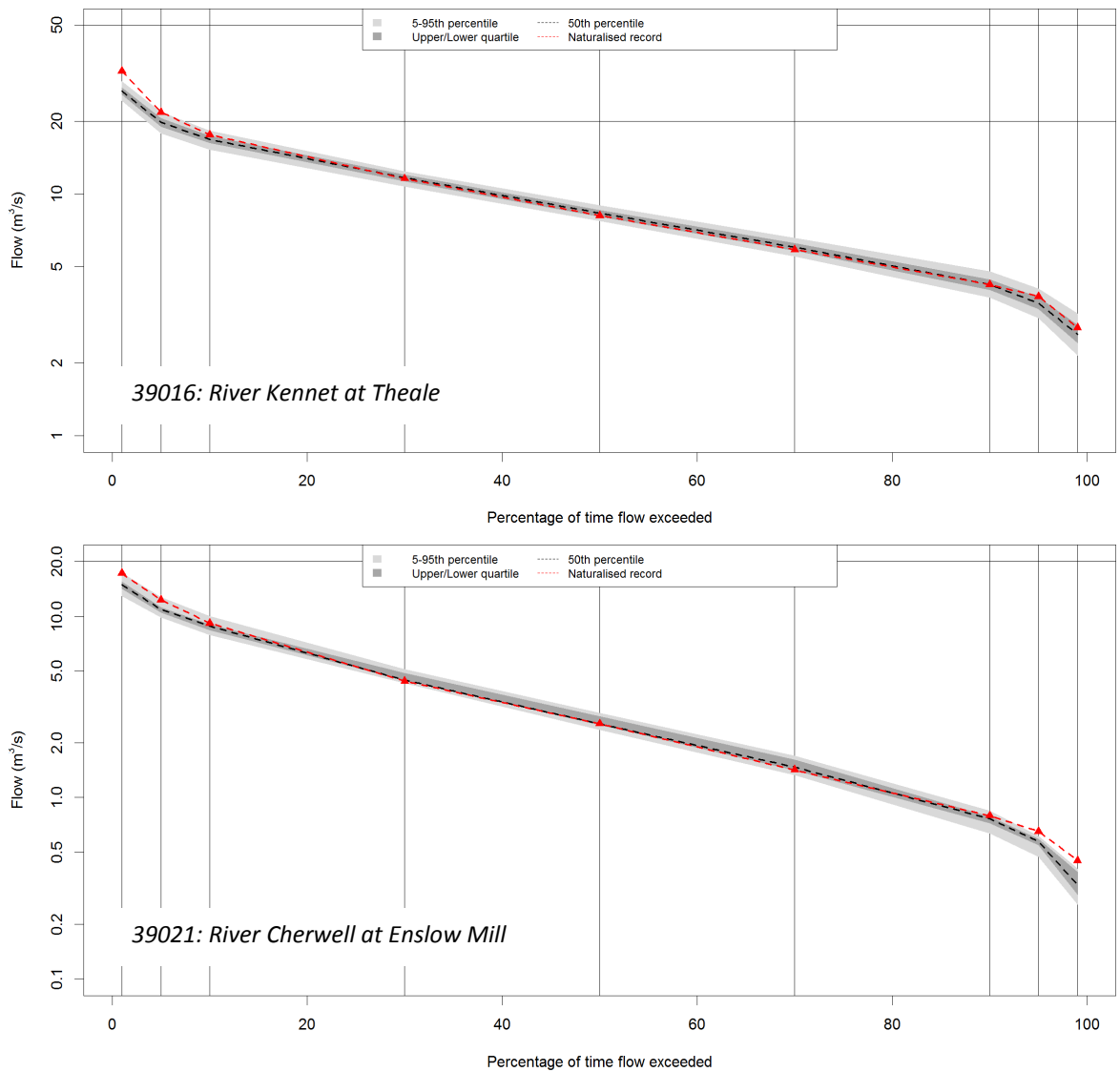
**Figure 4-18: Simulated hydrograph for the River Thames at Kingston during 1996 and 1997 shown in grey shading and black dashed line. Equivalent values from the historical record shown in red.**

#### 4.4.2 *Reproduction of flow metrics at other locations in the Thames basin*

As well as considering the performance of the ensemble for the River Thames at Kingston, the other locations used in developing the MSME have also been assessed (see Figure 3-1 for locations). Figure 4-19 and Figure 4-20 present the FDCs for the four locations other than Kingston in the Thames basin for the length of their respective historical records. The plots show the central modelled estimate, upper and lower quartiles and the 5<sup>th</sup> and 95<sup>th</sup> percentiles and demonstrate that, in general, the historical flow records appear to be well reproduced, with the largest divergences at low flows for catchments 39006 (Newbridge) and 39021 (Enslow Mill) and at high flows for catchments 39011 (Tilford) and 39016 (Theale). The relatively narrow prediction bands, particular for catchment 39011, may be a reflection of the relatively limited number of ensemble models available (see Table 4-4).

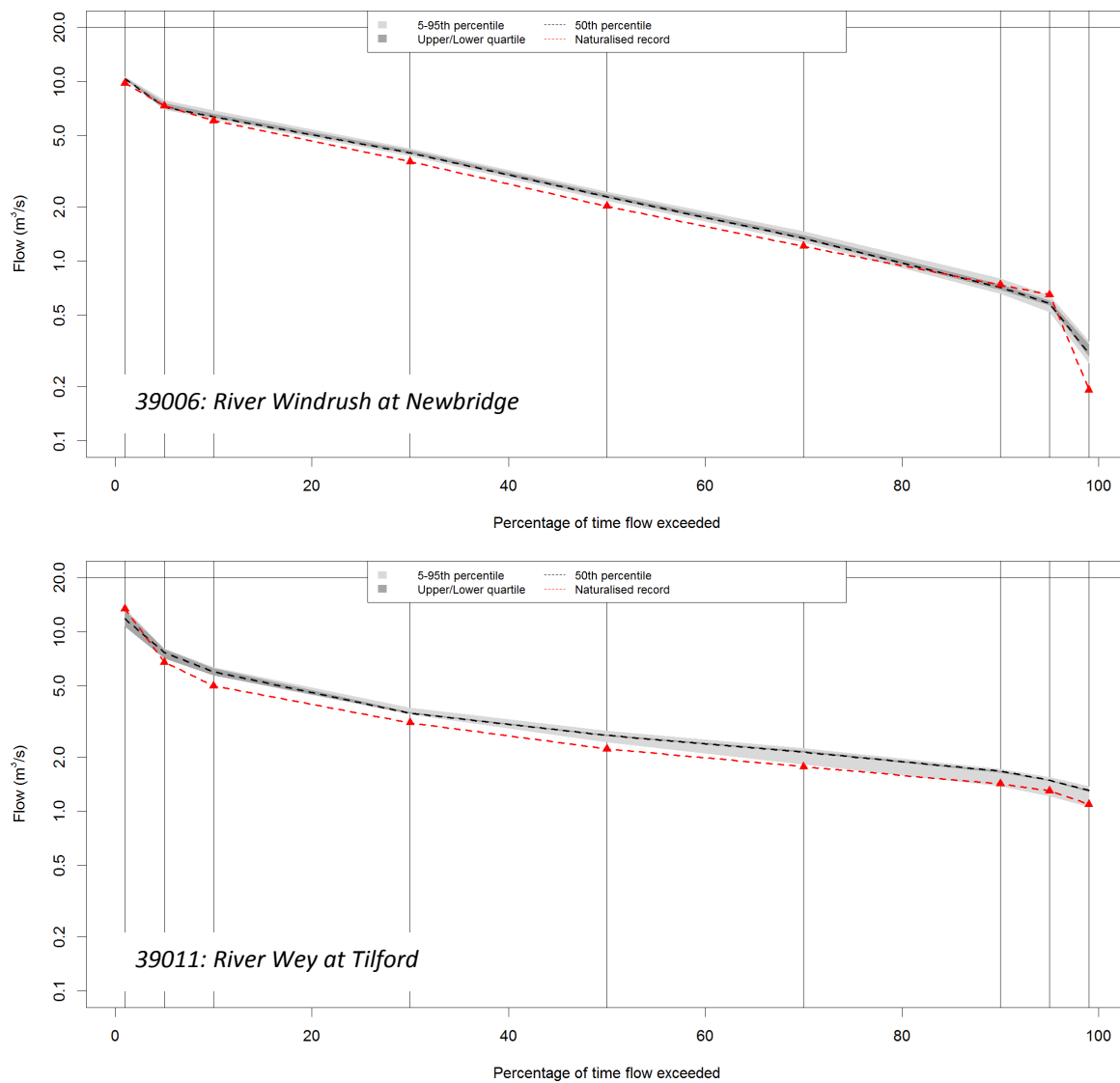


**Figure 4-19: Historical (red) and simulated (black/grey) FDCs for the period 1961 to 2008 for: top plot) River Windrush at Newbridge (NRFA catchment: 39006) and bottom plot) River Wey at Tilford (NRFA catchment: 39011).**

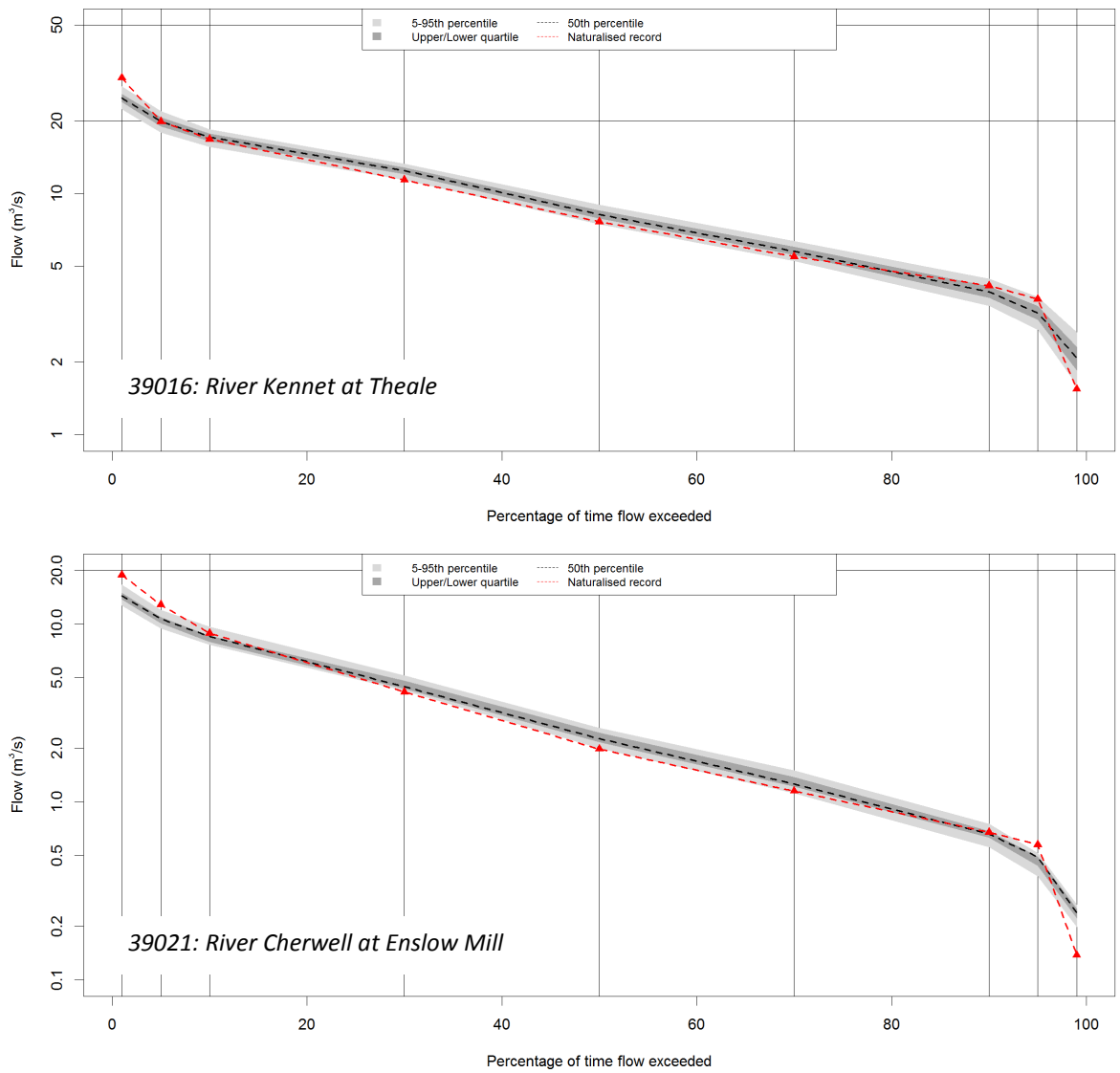


**Figure 4-20: Historical (red) and simulated (black/grey) FDCs for: top plot) River Kennet at Theale - from 1963 to 2008 (NRFA catchment 39016) and bottom plot) River Cherwell at Enslow Mill – from 1967 to 2008 (NRFA catchment 39021).**

Figure 4-21 and Figure 4-22 focus on the 1971 to 1980 period which contains the lowest mean summer flow sequence in the historical record (see Table 3-2). The plot shows the central modelled estimate, upper and lower quartiles and the 5<sup>th</sup>-95<sup>th</sup> percentiles and demonstrates that this section of the historical flow record is reasonably well reproduced, with the largest divergences for the Wey catchment at Tilford (39011) which is also the location for which the fewest models were retained. For all four sites, predictions at very high and very low flows demonstrate the poorest relative performance.



**Figure 4-21: Historical (red) and simulated (black/grey) FDCs for the 1971 to 1980 period for: top plot) River Windrush at Newbridge and bottom plot) River Wey at Tilford.**

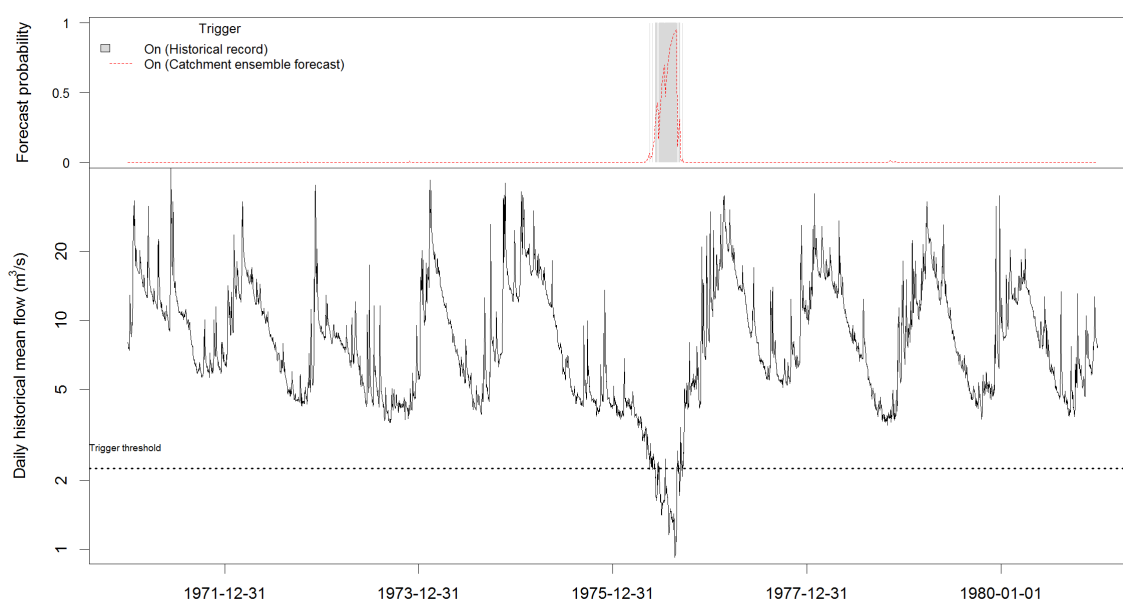


**Figure 4-22: Historical (red) and simulated (black/grey) FDCs for the 1971 to 1980 period for: top plot) River Kennet at Theale and bottom plot) River Cherwell at Enslow Mill.**

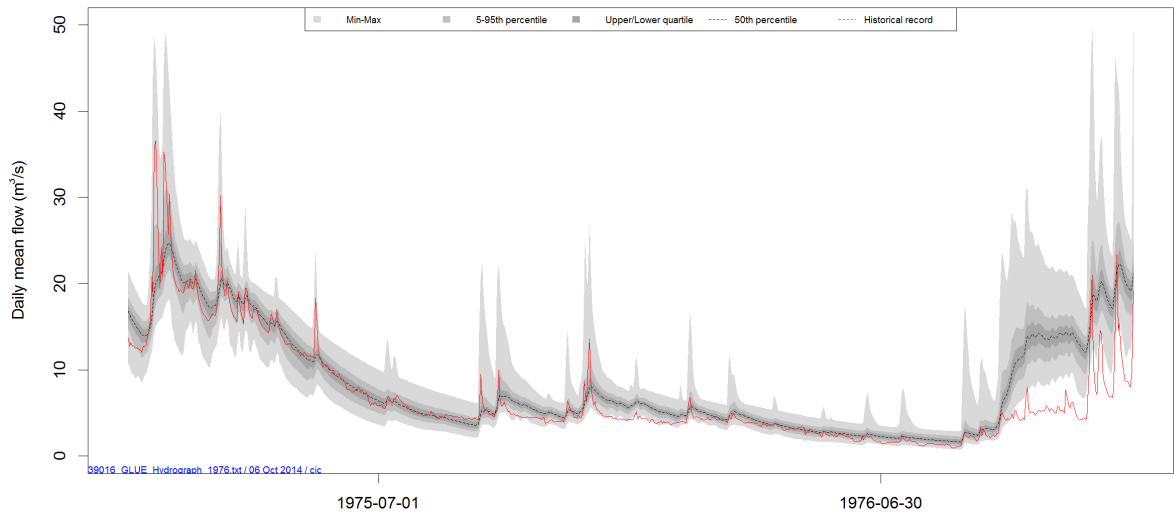
#### 4.4.2.1 Drought Plan trigger on River Kennet at Theale (39016)

The River Kennet at Theale is an important location for Thames Water's Kennet Valley water resource zone (which includes Reading and Newbury) and Thames Water's Drought Plan (2010) makes reference to a trigger threshold of 195 MI/d (equivalent to 2.3 m³/s) in the River Kennet at Theale (39016) for implementing drought response measures.

Figure 4-23 identifies such sequences in the historical and modelled flow series demonstrating only a single occurrence during the 1961 to 2008 historical record, in the summer of 1976. The plot presents the forecasting skill of the MSME which demonstrates reasonable reproduction of this sequence. The duration of the trigger threshold being met in the historical record is 91 days, and the RFAT and RFFAT scores (see Table 4-6) are 0.59 and 0.95 respectively. The historical and MSME predicted hydrographs for the 1975 to 1976 period are shown in Figure 4-24 and demonstrate similar characteristics to that exhibited for Kingston in terms of over-prediction during the recovery of autumn 1976, and to a lesser degree, during the modest recovery during the autumn of 1975.



**Figure 4-23: Reproduction of trigger threshold activations (195 MI/d, ~2.3 m<sup>3</sup>/s, for the River Kennet at Theale) during Jan 1971 to Dec 1980 shown as a red line in top plot. Grey shading indicates periods in historical record where trigger threshold is crossed, with historical flow record itself shown in black on bottom plot.**



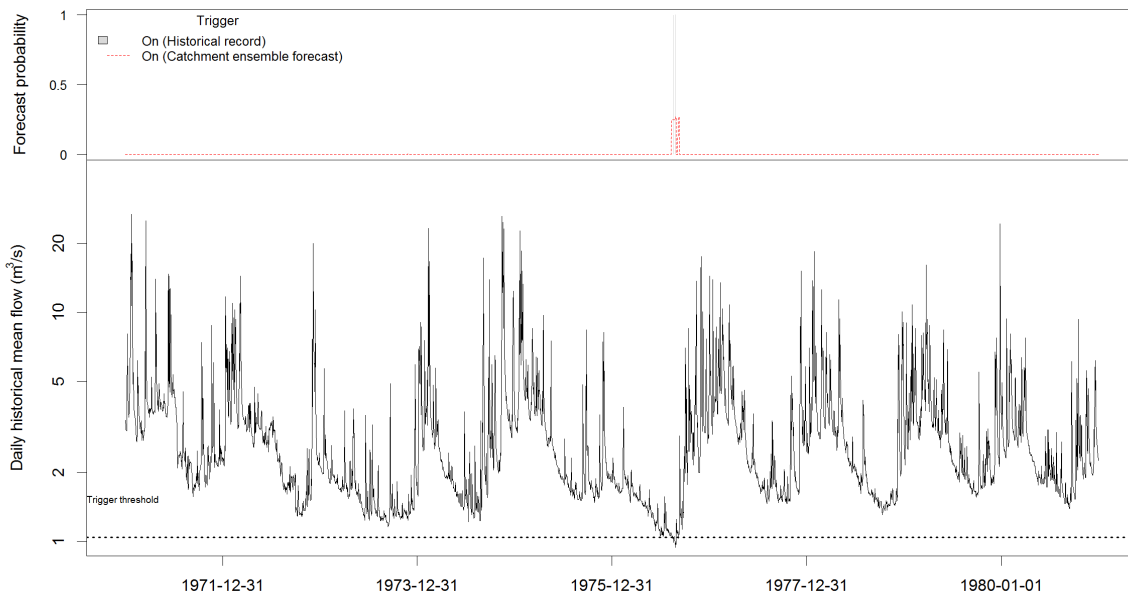
**Figure 4-24: Reproduction of the flow hydrograph for the River Kennet at Theale during 1975 and 1976. Simulated flows shown as grey shading (5<sup>th</sup> to 95<sup>th</sup> percentile estimate) and black dashed line (50<sup>th</sup> percentile estimate). Equivalent values from the historical record shown in red.**

#### 4.4.2.2 Drought Plan trigger on River Wey at Tilford (39011)

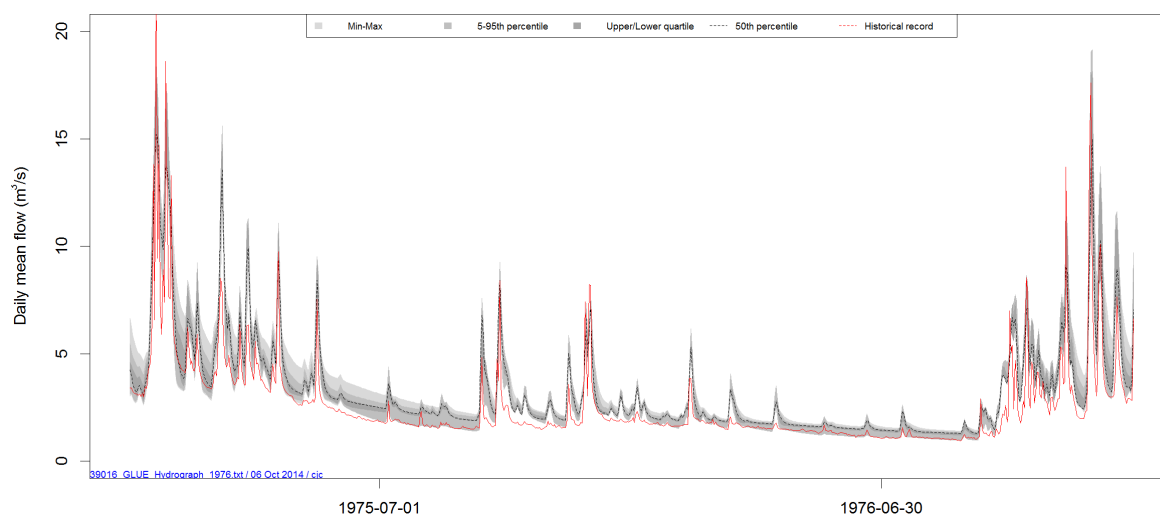
The River Wey at Tilford is a key location for Thames Water's Guildford water resource zone and Thames Water's Drought Plan (2010) makes reference to a flow threshold of 90 MI/d (1.0 m<sup>3</sup>/s) in the River Wey at Tilford (39011) that, when experienced for 5-days or longer, may trigger drought response measures – with no such sequences are present in the 1961 to 2008 record. Figure 4-25 shows the 1971 to 1980 period within which the threshold is met but not for the minimum 5 days and is the closest the historical record gets to satisfying this trigger condition. The MSME shows that some models suggest the conditions are met but the overall forecast probability is significantly below 0.5.

The hydrograph for this period is shown in Figure 4-26 and demonstrates similar traits to that shown for 39001 (Kingston) and 39016 (Theale), noting the fewer number of models used in prediction which may be reflected in the narrower uncertainty bands.





**Figure 4-25: Reproduction of trigger threshold activations (90 MI/d,  $\sim 1.0 \text{ m}^3/\text{s}$ ) for the River Wey at Tilford) during Jan 1971 to Dec 1980 period shown as a red line in top plot. Grey shading indicates periods in historical record where trigger threshold is crossed, with historical flow record itself shown in black on bottom plot.**



**Figure 4-26: Reproduction of the flow hydrograph for the River Wey at Tilford during 1975 and 1976. Simulated flows shown as grey shading (5<sup>th</sup> to 95<sup>th</sup> percentile estimates) and black dashed line (50<sup>th</sup> percentile estimate). Values from the historical record are shown in red.**

## 4.5 Chapter summary

This chapter has described the development of a multi-site hydrological model ensemble (MSME) to provide spatially coherent river flows at multiple sites of interest for water resources planning in the Thames basin. The development of the MSME has considered two hydrological model structures, two alternative methods for estimating PET, six performance criteria and five time-periods to evaluate model performance as part of a Generalised Likelihood Uncertainty Estimation (GLUE) methodology (Beven & Binley, 1992).

In developing the MSME, this research has suggested that the PDM-AQ model structure is a more appropriate model structure for the Thames basin than the Thames Catchment Model (TCM) structure which has been the most commonly used model structure on the Thames catchment in previous industry and academic water resources studies (e.g. Thames Water, 2014; Wilby & Harris, 2006; Environment Agency, 2009; Manning et al., 2009; Borgomeo et al., 2014; Walsh et al., 2015). Performance of both model structures at interior locations demonstrated particular challenges (in identifying acceptable models) at some study sites (i.e. River Windrush at Newbridge and the River Wey at Tilford) and resulted in all TCM based models being rejected and not taken forward as part of the final ensemble.

However, the relative performance of the model structures will be partly dependent upon the design of the GLUE based approach to considering uncertainty, including the choice of performance metrics and conditioning periods. It is worth noting that each model structure, with the four HRUs used, has 20 'free' parameters and further research could include a systematic diagnostic evaluation of the model structures to better understand why certain characteristics of the hydrological signature are particularly difficult to reproduce.

Development of the MSME demonstrated that whilst alternative methods for estimating PET do lead to differences in the overall water balance, the model structures used in this research were able to compensate for such differences in their parameterisation. Consequently, models driven by both methods for estimating PET have been incorporated in the final ensemble. The development also suggested that using a gridded soil moisture accounting component, driven by gridded climatic data, within the selected model structures does not necessarily lead to improved model performance compared to using a semi-lumped forms and catchment average climatology. It might be that in catchments with greater spatial variation in hydrogeology, topography and precipitation pattern benefits maybe more evident. Consequently, this form of the model structures was not taken forward as part of the MSME.

The remainder of this chapter has evaluated the ensembles' skill in reproducing a range of water resources relevant hydrological characteristics to identify its potential strengths and weaknesses and to guide the interpretation of subsequent predictions in later chapters. This evaluation highlighted that flow duration curves were reasonably well reproduced, particularly at Kingston, apart from for the highest flows (significant divergence with the historical record at Q10 and above was apparent). Performance at interior locations demonstrated that, in general, the historical flow records across the study sites were generally well reproduced by the MSME, with the largest divergences at low flows for Newbridge (39006) and Enslow Mill (39021) and high flows for Tilford (39011) and Theale (39016).

In terms of dry years, of particular interest to water resources planners, the ranking of years in the historical record (as measured by mean summer flow) was accurately reproduced at Kingston by the MSME but there was a tendency to over-estimate the mean summer flow values. The exception was for the lowest mean summer flow from the most recent evaluation period (2005). Examination of the associated flow hydrographs demonstrated that whilst the MSME was able to reproduce flows during the summer of 1976 reasonably well, including the timing of relatively small peak flows from short duration rainfall events when soil moisture deficits are high, it was not able to accurately reproduce the wetting-up period in the autumn of 1976.

Furthermore, the MSME tended to sustain flows after each of the modest summer rainfall events for longer than was evident in the historical record. In general, it appears that the MSME 'released' more water to the river during, and just after, periods of high soil moisture deficits and it is speculated that during such periods more water may be taken up to replenish soil moisture and low aquifer levels (e.g. Oni et al., 2016). However, the exception to this, observed during the most recent evaluation period (2001 to 2008), may have been influenced by the MSME's consistent underestimation of flows throughout this period.

The challenges in reproducing the physical processes during the termination of droughts has been highlighted by Parry et al. (2016b) which noted that the extent to which climate and hydrological models are able to simulate observed drought termination events remains an open question. Furthermore, these periods of sustained low flow are likely to be subject to significant management and regulation of the river to maintain flows for navigation, and that it might be also expected that upstream abstractors may be attempting to capture such flows to mitigate supply problems they may have experienced during the previous period of low flow. Parry et al. (2016a), as part of a systematic review of drought termination, reported that abstractions from surface and groundwater sources during drought development may artificially extend the duration of the

drought termination phase. Therefore, there is also likely to be relatively greater uncertainty associated with flow records during such periods when artificial influences are also likely to have the largest relative impact on river flows.

When looking at the conditions that trigger the operation of strategic schemes to provide additional support during droughts, this research highlights the differences in predictive accuracy of the MSME during different time-periods. During the 1971 to 1980 and 1981 to 1992 periods, the overall frequency of trigger conditions being met at Kingston predicted by the MSME was close to the number in the historical record but there were several false positives and false negatives. Looking at the hydrographs it is apparent that the inaccuracies are related to those periods where the historical flow was close to the trigger threshold rather than the more extreme (both in magnitude and duration) sequences in the historical record.

Considering the reproduction of winter flows, the MSME broadly reproduced the relative ranking of years reasonably well but particular inconsistencies compared to the historical record during the winter of 1996/1997 and 2004/2005 and, to a lesser extent, 1975/1976 were evident. The summer of 1996 (and 1995 to a lesser extent) was particularly dry so this may be related to the 'wetting-up' problem previously identified. The performance values used to provide the conditional probabilities are derived using a range of measures which include criteria that specifically target lower flows such as the drought-error metric and therefore may be less appropriate for considering high flow characteristics.

The issues identified above, along with the difficulties faced using the TCM model structure, highlights that further research into the strengths and weaknesses of different model structures is required for water resources planning, particularly in the representation of key processes during droughts. The range of model predictions, conditional on the criteria used to measure performance, presented by the model ensemble also highlights the degree of sensitivity of model predictions to the uncertainty associated with model parameterisation. Finally, a thorough review of the data records for the Thames basin is warranted to understand the degree to which inconsistencies between the MSME predictions and the historical record are related to inaccuracies in the flow record rather than limitations in the model ensemble.

The MSME developed as part of this chapter is taken forward to explore the potential impacts of climate change on river flows across the Thames basin and the relative sensitivity of such predictions to different sources of uncertainty in Chapter 5. These river flow predictions, and the associated sources of uncertainty, are then examined in the context of the projected performance of London's water resource system in Chapter 6.

## Chapter 5 Impacts of climate change on river flows for water resources planning

### 5.1 Introduction

As presented in Chapter 2, it is anticipated that climate change is likely to have a significant impact on water resources across the United Kingdom (e.g. Wilby & Harris, 2006; New et al., 2007a; Fowler et al., 2008; Cloke et al., 2010; Christerson et al., 2012; Borgomeo et al., 2014) and any significant reduction in water availability, particularly in areas of the country considered to be already “water stressed” (Environment Agency, 2008), could have major implications for our quality of life and the environment.

This chapter examines the two alternative sources of climate projections recommended for water resources planning in the UK. These climate projections are the UKCP09 ‘probabilistic’ projections (Murphy et al., 2009a – see Section 2.2.2.1) and the eleven-member Future Flows transient climate projections (Prudhomme et al., 2012 – see Section 2.2.2.2).

### 5.2 Objectives

In this chapter the following research questions (see Section 1.2) are investigated:

- RQ2.** How can the climate projections currently recommended for use by water resource planners in the UK be used as part of an ensemble modelling study and how appropriate are these projections for water resources planning? How significant is the method used to estimate PET as part of a climate change impact assessment?
- RQ5.** What is the projected magnitude of the impacts of climate change on river flows across the Thames catchment and what is the trajectory of these changes beyond the industry standard 25 year planning horizon?

In addressing these questions, Chapter 5 examines the changes to precipitation and PET (the two input variables typically used to drive hydrological models) projected by the UKCP09 and Future Flows climate products. The Multi-Site Model Ensemble (MSME) described in Section 4.3.4 is then used to assess the impacts of climate change on river flows across the study locations (see Figure 3-1) for the 2030s, 2050s and 2080s. As part of this assessment the impacts from using the two methods to calculate PET described in Section 3.4.2.1 are contrasted and the relative magnitude of climate modelling uncertainty is compared to that associated with hydrological modelling uncertainty.

## 5.3 Climate change projections for water resource planners

### 5.3.1 *The 'probabilistic' UKCP09 climate change projections*

The UK Climate Projections (UKCP09), described in Section 2.2.2.1, represent the fifth generation of climate scenarios that have been produced for the UK since the publication of the first scenarios (CCIRG91) in 1991 (see Figure 2-1).

Figure 5-1, Figure 5-2 and Figure 5-3 present seasonal changes in mean daily precipitation and temperature in the Thames basin for the full UKCP09 10,000 member ensemble for the 2030s, 2050s and 2080s respectively. These plots demonstrate the large uncertainty associated with the UKCP09 projection. Under the 2080s Medium Emission scenarios the range of predicted change in summer precipitation for the Thames basin varies from -54% (5<sup>th</sup> percentile) to +17% (95<sup>th</sup> percentile), with a central estimate (50<sup>th</sup> percentile) of -22%. For winter precipitation, the 5<sup>th</sup>, 50<sup>th</sup> and 95<sup>th</sup> percentiles estimates are -2%, +20% and +57% respectively. For the change in summer temperature, the 5th, 50th and 95th percentiles are +2, +4 and +7 degrees Celsius respectively. These wide ranges highlight the high level of uncertainty associated with the UKCP09 climate projections.

#### 5.3.1.1 Sampling of the UKCP09 climate change projections

The UKCP09 user-guidance recommends that each member from the 10,000 member ensembles are considered equally likely (Murphy et al., 2009b). To reduce the computational resource required for the hydrological modelling presented in this thesis, a representative sample for each time-horizon was required for each time-horizon.

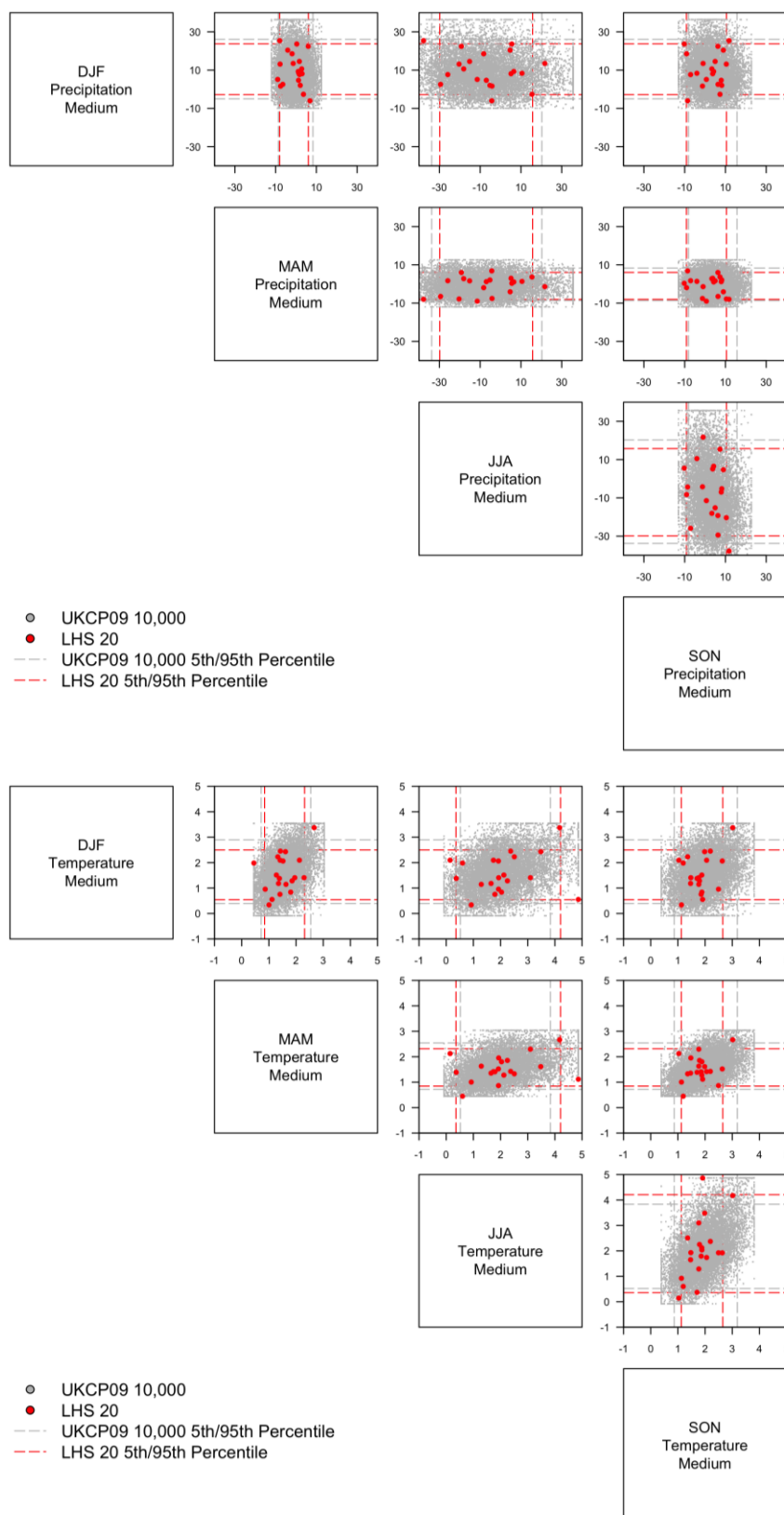
In their study of six UK catchments, Charlton & Arnell (2014) suggested that whilst around 20 randomly sampled UKCP09 scenarios can provide a reasonably robust estimate of the range in potential changes, a larger number is required to derive a robust estimate of the broad shape of the distribution of changes. As an alternative to a random sampling approach Christerson et al. (2012) adopted Latin Hypercube Sampling (McKay et al., 1979) in their study of 70 UK catchment using the UKCP09 projections. In this research, the 10,000 UKCP09 projections were split into multiple 'blocks' across eight dimensions (seasonal changes in precipitation and temperature) with the LHS used to randomly select the blocks from within which a sample of 20 was randomly taken. Each of these sampled scenarios was then considered equally likely in their representation of the full UKCP09 ensemble. For the Thames catchment Christerson et al. (2012) reported that a sample size of 20 was sufficient to capture most of the uncertainty in terms of flow impacts on the Thames.

In this thesis this LHS approach has also been applied to sample 20 representative scenarios for the 2030s, 2050s and 2080s epochs, all under the Medium Emission scenario. As outlined above, this reduced sample is intended to broadly reflect the uncertainty present in the full 10,000 member ensembles and the efficacy of this sampling approach is examined in Figure 5-1, Figure 5-2 and Figure 5-3 which present seasonal changes in mean daily precipitation and temperature for each set of 20 sampled scenarios alongside those for the full UKCP09 10,000 member ensemble. These figures compare the range and coverage of the samples with that of the corresponding 10,000 member ensembles and demonstrate that the 5<sup>th</sup> and 95<sup>th</sup> percentiles from the sub-samples are similar to those from the full ensemble. Further analysis of the sampling efficacy is explored in the following section.

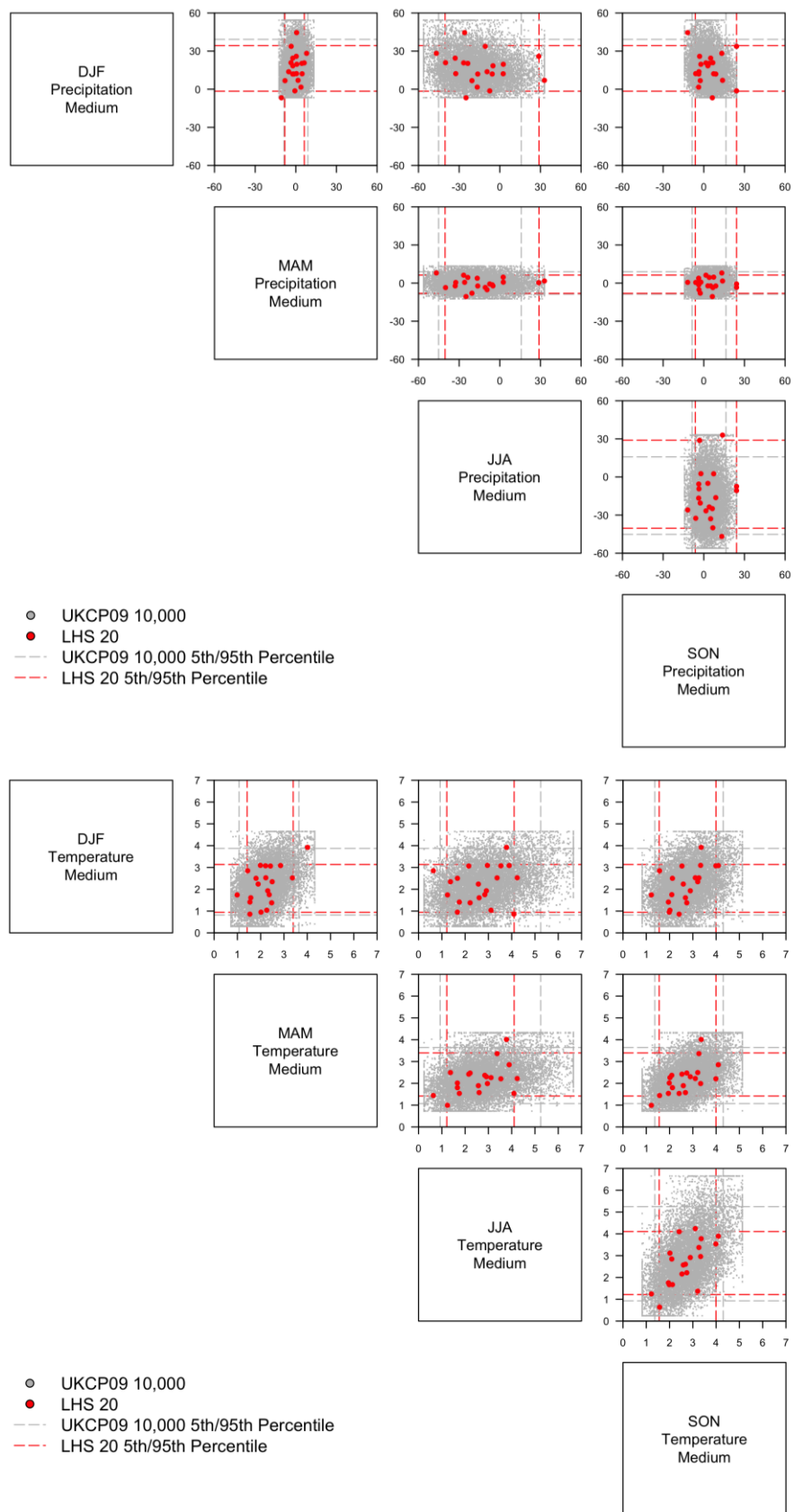
#### 5.3.1.2 Estimation of Potential Evapotranspiration (PET)

Previous research (e.g. Kay and Davies, 2008; Bell et al., 2011; Prudhomme & Williamson, 2013; Kingston et al., 2009; Christerson et al., 2012) has highlighted that different methods for estimating PET, and in particular calculating the changes in PET under a future climate, can vary significantly which presents a challenge to the water resources planner, particularly as the impact of such changes may be particularly significant at low flows (Charlton & Arnell 2014).

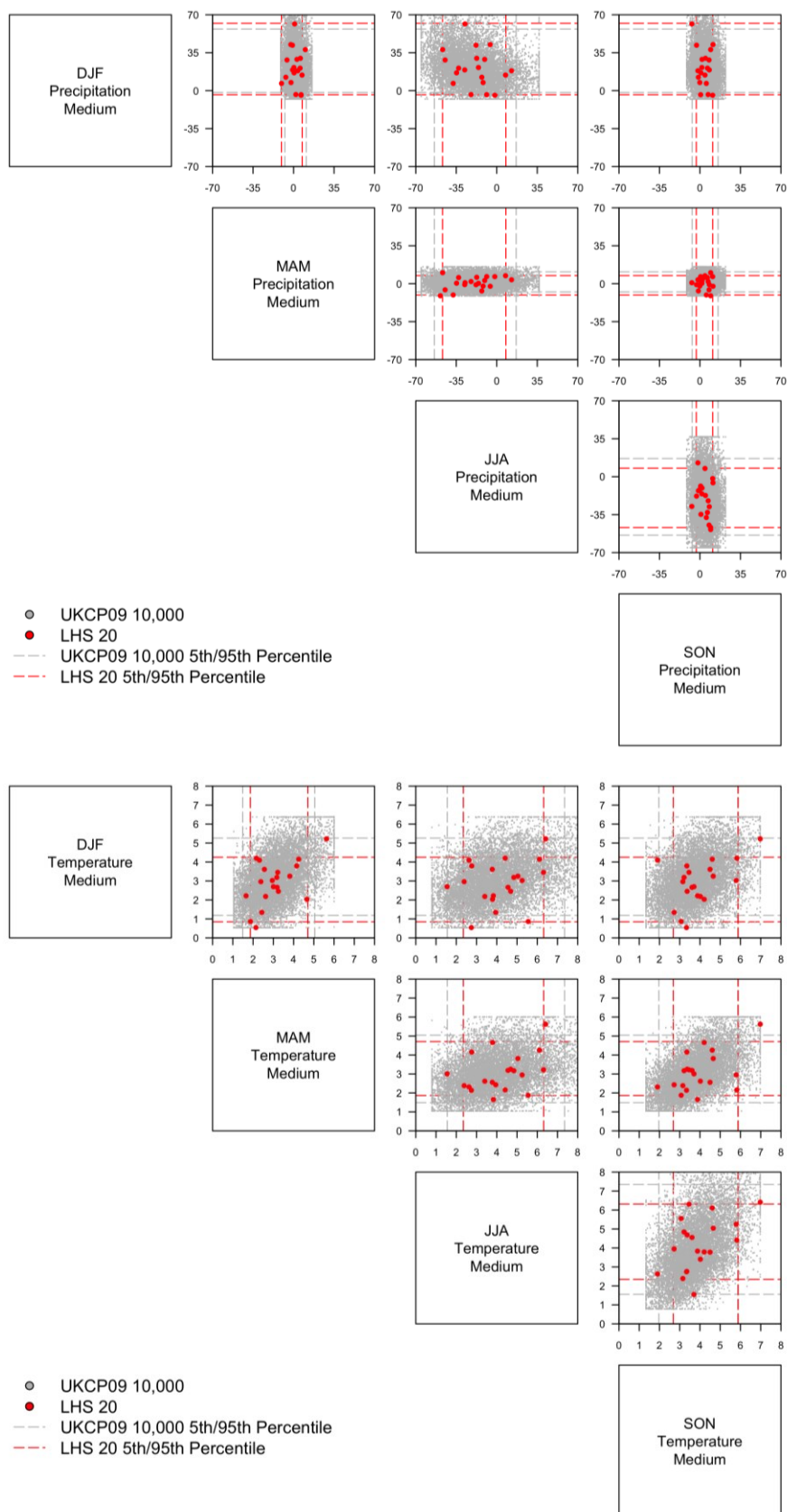




**Figure 5-1: Seasonal change (percentage) in mean daily precipitation and temperature (in degrees Celsius) from the UKCP09 2030 Medium Emission scenario for the full 10,000 member ensemble and the 20-member sub-sample. Changes are relative to a 1961 to 1990 baseline period.**



**Figure 5-2: Seasonal change (percentage) in mean daily precipitation and temperature (in degrees Celsius) from the UKCP09 2050 Medium Emission scenario for the full 10,000 member ensemble and the 20-member sub-sample. Changes are relative to a 1961 to 1990 baseline period.**



**Figure 5-3: Seasonal change (percentage) in mean daily precipitation and temperature (in degrees Celsius) from the UKCP09 2080 Medium Emission scenario for the full 10,000 member ensemble and the 20-member sub-sample. Changes are relative to a 1961 to 1990 baseline period.**

In this thesis, two methods of estimating PET are considered (see Section 3.4.2.1). The Oudin (2005) method uses mean daily temperature as the input climate variable whilst the modified Penman-Monteith (Prudhomme & Williamson, 2013) method uses mean daily relative humidity and cloud cover as well as temperature – noting that it is assumed that wind speed remains constant (See Section 2.2.2.1). A similar assumption with regards to wind speed was adopted by Lopez et al. (2009) and Charlton & Arnell (2014) in their studies of climate change impacts on water resources system performance and river flows respectively (see Section 2.3.2).

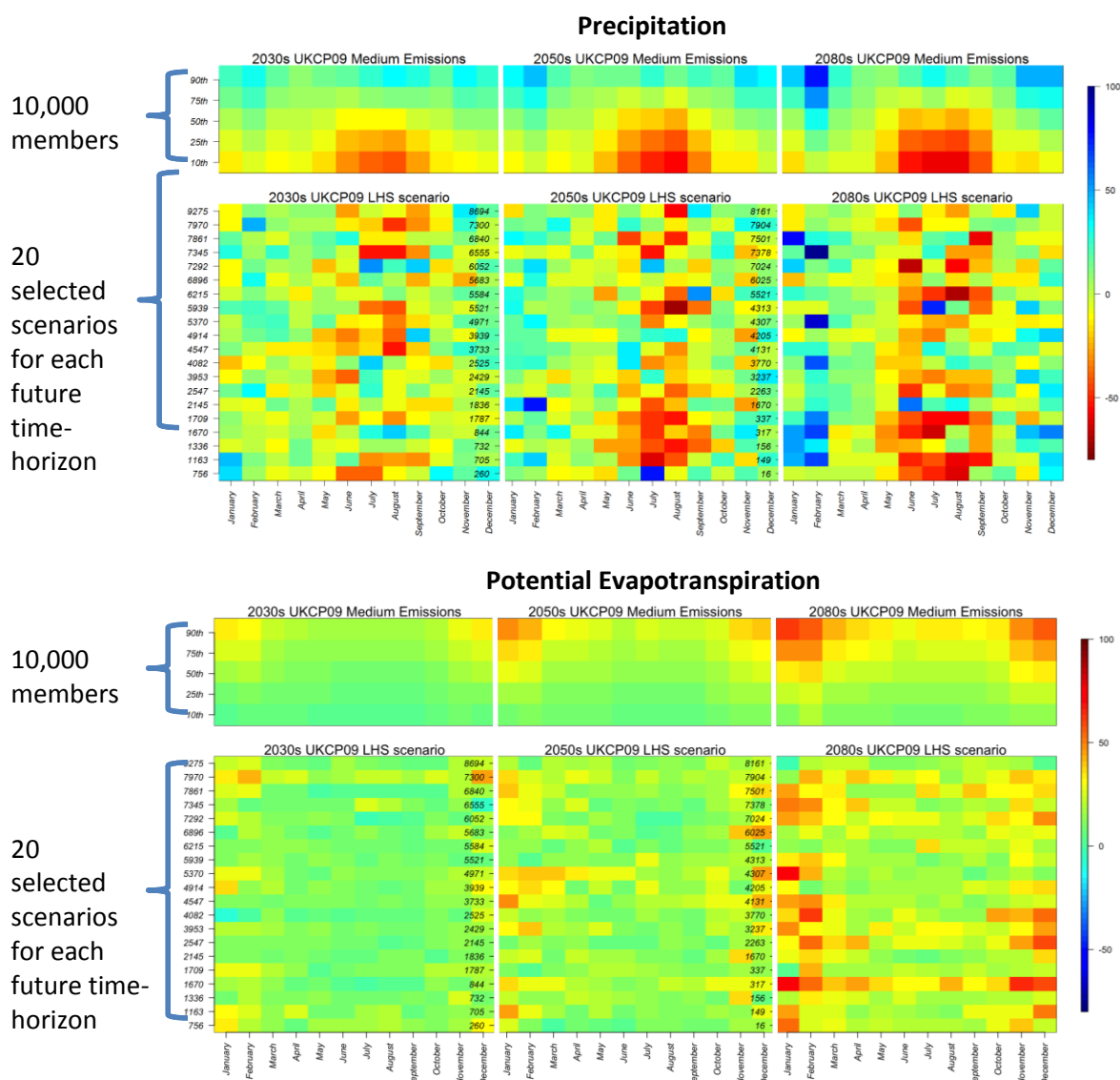
With regards to wind speed, changes can be derived from the underlying Regional Climate Models (RCMs) and previous research (UKCP09, 2017) has shown these changes are small –predominantly changes of up to -3% reported by the 2080s (from a 1961 to 1990 baseline) for most of the UK, consistent with evidence from outputs from GCMs (Murphy et al., 2009b). Further discussion of the application of the two methods for estimating PET is presented in Section 3.4.2.1.

Figure 5-4 presents monthly changes in PET (using the Oudin method) as well as precipitation for selected percentiles for the full 10,000 member ensembles and for each scenario from the 20-member sub-samples. These plots demonstrate that the reduced samples for each future time-horizon provide reasonable coverage of the uncertainty associated with precipitation and PET present within the full ensembles for each time-horizon. However, the efficacy of the sampling towards the extremities (for example temperature changes in the autumn and winter in the 2080s) should be recognised and may result in conservative upper and lower bound estimates of the projected impacts of climate change, albeit that the 5<sup>th</sup> and 95<sup>th</sup> percentiles from the sub-samples are similar in value to those from the full 10,000 member ensemble (see Figure 5-1, Figure 5-2 and Figure 5-3).

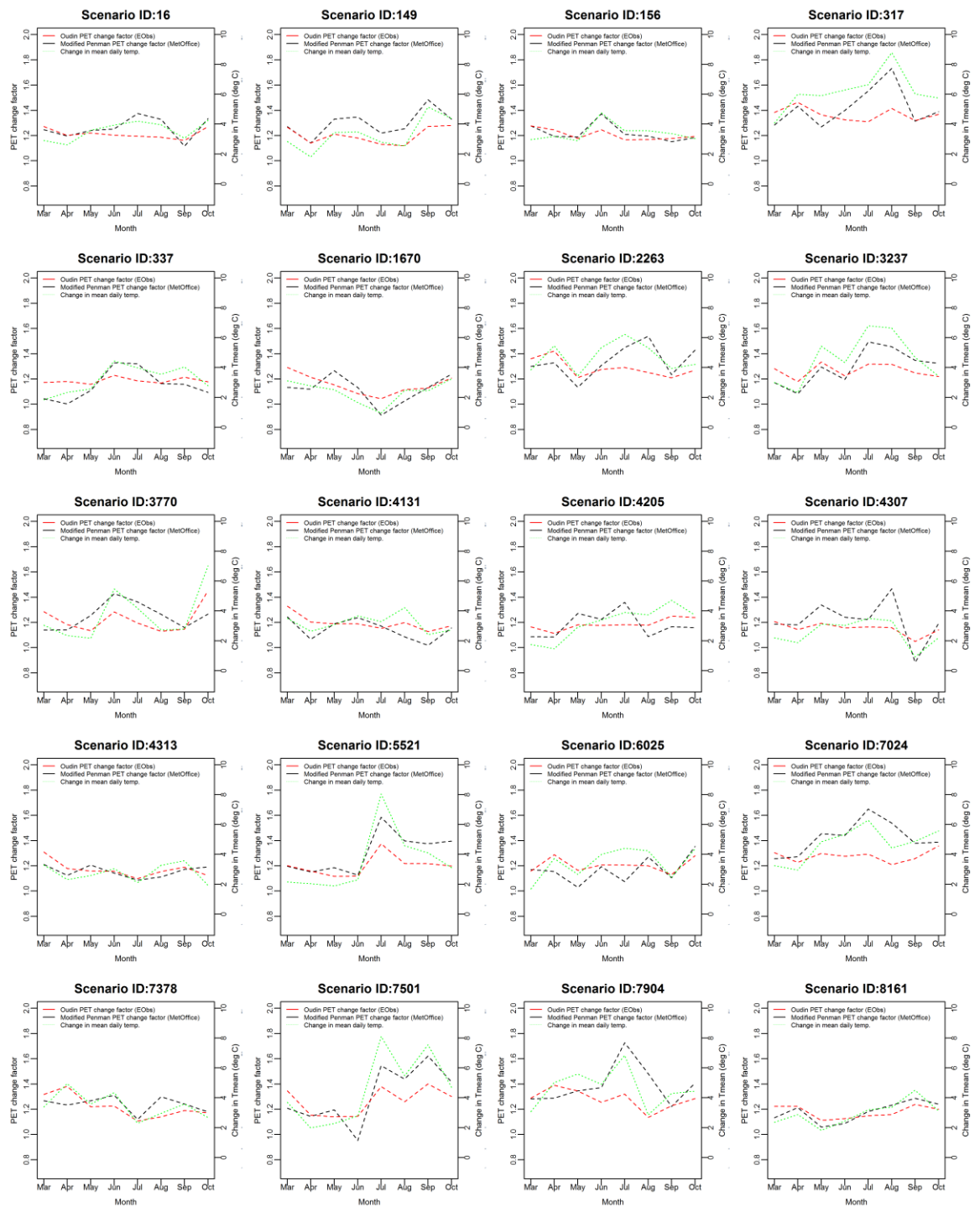
Figure 5-5 presents the projected changes in mean monthly PET, along with the projected changes in mean daily temperature, for each sample for the 2080s (with similar plots for the 2030s and 2050s included in Appendix B) – note that the months from November through to February are not shown because these are considered less important and include high percentage change factors due to the relatively small absolute values of PET during these months.

PET can be seen to increase through the future time-horizons and the plots highlight that for scenarios that include large changes in temperature (e.g. change in temperature in August under the 2080s *scenario 2263*) the modified Penman-Monteith method typically projects significantly higher changes in mean monthly PET than the changes calculated using the Oudin formula. The variability across the twenty UKCP09 scenarios, for the 2080s is summarised in Figure 5-6 in the form of a box plot of the variability in monthly PET, from March to October for both methods with

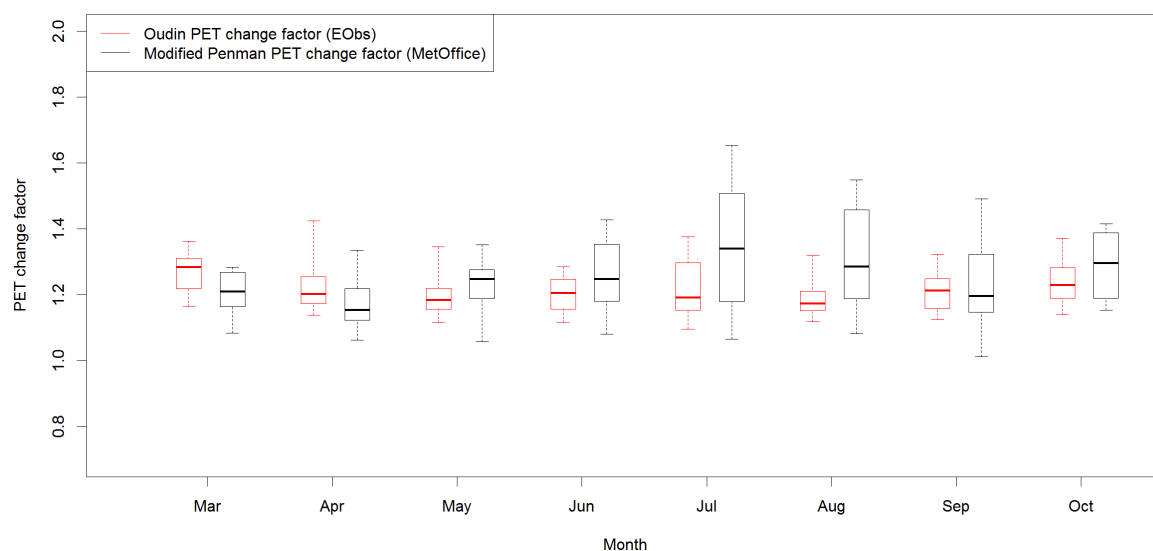
the modified Penman-Monteith method demonstrating much greater variation across the scenarios, particularly in the summer months, as well as a tendency to project greater changes overall.



**Figure 5-4: Mean monthly change (as a percentage) in precipitation (top plot) and PET (bottom plot, using the Oudin method) from the UKCP09 2030, 2050 and 2080 Medium Emission scenarios. Values shown are for the 10<sup>th</sup>, 25<sup>th</sup>, 50<sup>th</sup>, 75<sup>th</sup> and 90<sup>th</sup> percentiles for the full 10,000 member ensembles and each selected scenario from the 20-member sub-samples. Changes are relative to a 1961 to 1990 baseline. Each y-axis provides the UKCP09 scenario ID for each sampled scenario for each future time-horizon.**



**Figure 5-5: Comparison of mean monthly changes in PET under the Oudin (red dashed lines) and modified Penman-Monteith (black dashed lines) methods and the mean monthly change in mean daily temperature (green dashed lines) for 20 UKCP09 sampled projections for the 2080s under a Medium Emission scenario. Changes are relative to a 1961 to 1990 baseline.**



**Figure 5-6: Comparison of the distribution of mean monthly changes in PET under the Oudin (red) and modified Penman-Monteith (black) methods for 20 UKCP09 sampled projections for the 2080s under a Medium Emission scenario. Boxplot whiskers extend to the 5<sup>th</sup> and 95<sup>th</sup> percentiles of the 20 sampled scenarios. Changes are relative to a 1961 to 1990 baseline.**

Prudhomme & Williamson (2013) compared estimated changes in PET using twelve different PET methods, including the two considered here, for a single transient (bias-corrected for temperature) output from a Regional Climate Model. The results presented by Prudhomme & Williamson (2013) are consistent with those presented here in that the Penman-Monteith suggested higher summer changes in PET than the Oudin method for the Thames area. Kay & Davies (2008) also reported that the Oudin and modified Penman-Monteith approaches led to marked differences in future changes in estimated PET by the end of the 21st century using outputs from five Global Climate Models and eight Regional Climate Models. Similarly, Ekström (2007) showed large differences between temperature driven and Penman-Monteith formulations in the North-West of England.

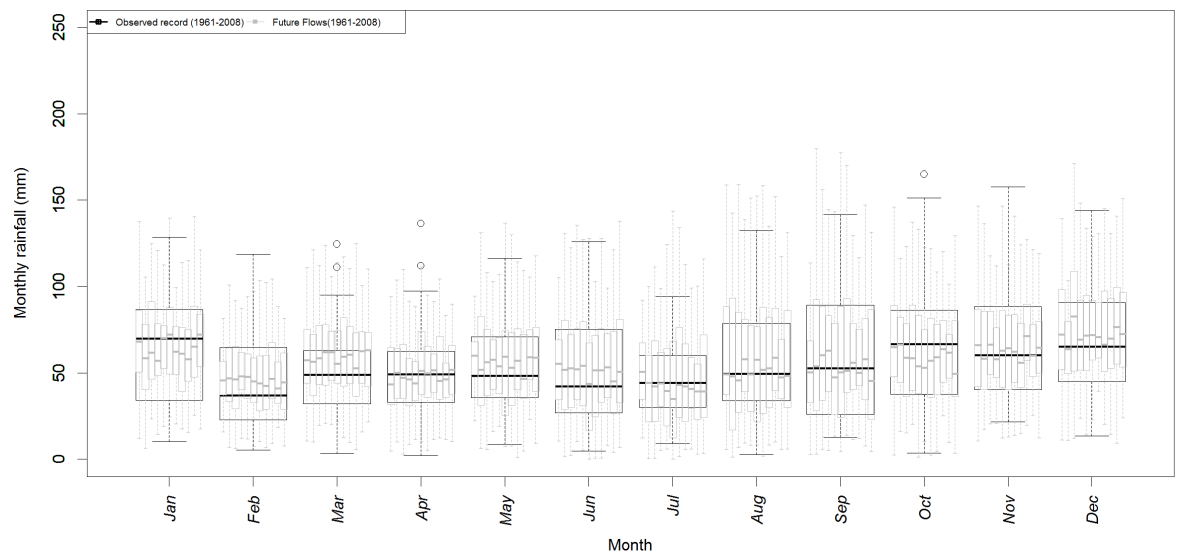
The Oudin formula is an empirical temperature based method and this may suggest that it may be considered less suitable under more extreme future climates as it can only take into account changes to a single variable. However, that could also be considered an advantage in that it does not require the input of a range of weather variables that may be less reliably predicted by climate models (Prudhomme & Williamson, 2013). The impacts of using these contrasting methods of estimating PET on future river flows (section 5.4.2) and water resources planning (section 6.4.2.2) are explored later in this chapter.

### 5.3.2 The transient Future Flows climate projections

As described in Section 2.2.2.2, the Future Flows (FF) climate projections (Prudhomme et al., 2012) provide an eleven-member ensemble of transient climate projections (from 1950 to 2098) specifically developed for hydrological and hydrogeological application in Great Britain. The following section compares the climatology present in the historical record over the 1961 to 2008 period with the same periods in the FF ensemble.

#### 5.3.2.1 Comparison of monthly rainfall and PET values in the FF ensemble with the historical record

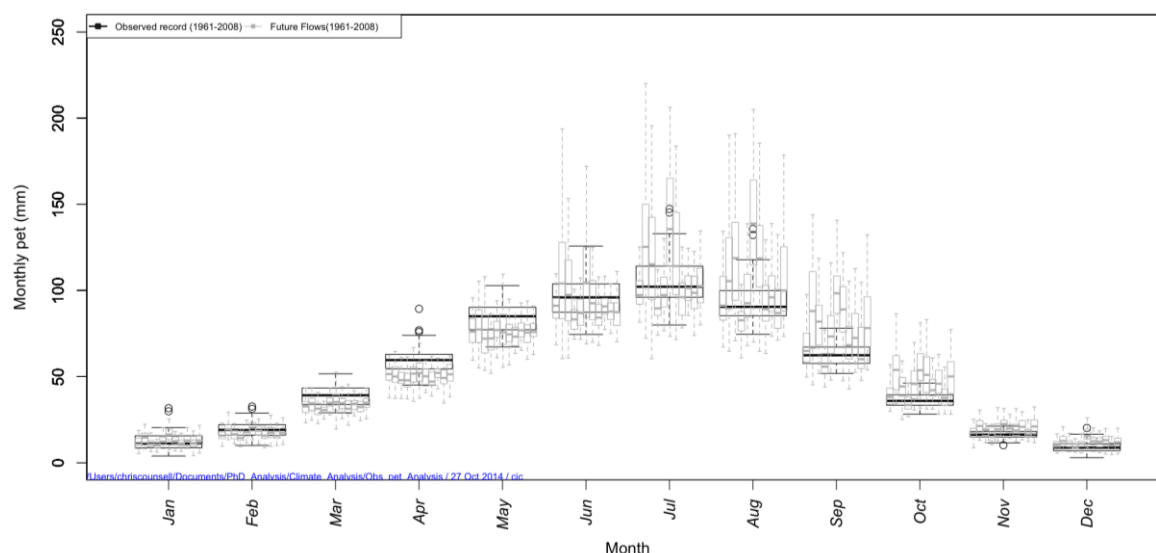
Figure 5-7 presents box-plots of the monthly precipitation totals from the historical record during the baseline period (1961 – 2008) and for the corresponding period for each FF ensemble member (note that the small circles are individual values that are more than 1.5 times the inter-quartile range from the extent of the inter-quartile range). The plots indicate that the range and magnitude of the monthly totals in the FF scenarios during this baseline period are broadly consistent with the historical record which would be expected given the monthly quantile-mapping approach taken to bias-correct precipitation in the development of the Future Flows scenarios (Newton et al., 2012).



**Figure 5-7: Box plots of monthly precipitation totals across the Thames basin from the historical record (E-Obs v 3.0 – see 3.4.2) as the wider black box-plots and the eleven-member Future Flows (FF) ensemble over the 1961 to 2008 time period (with each individual scenario shown as a light grey box plot).**

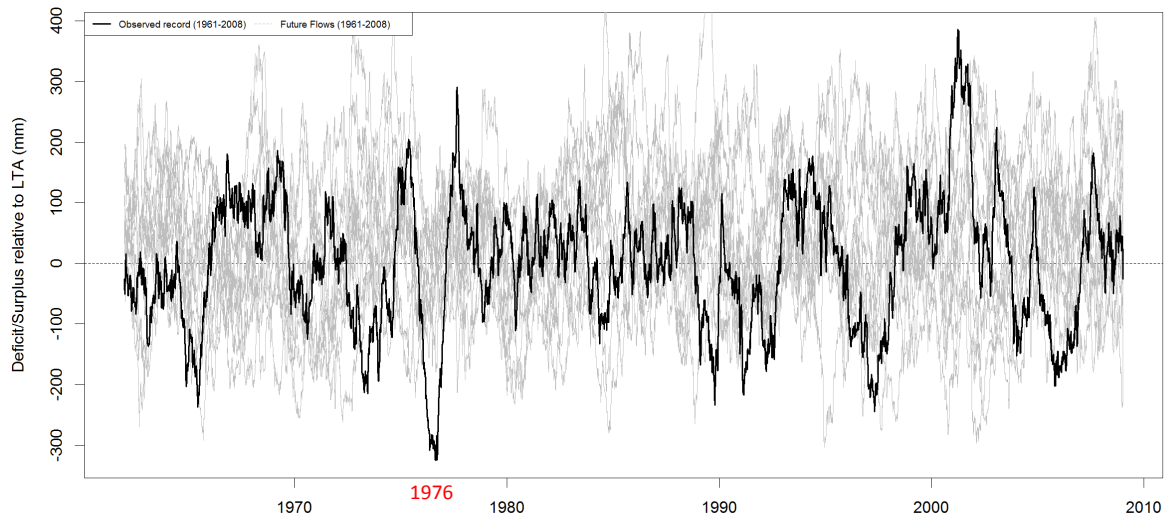


Figure 5-8 presents the same information for PET. For the FF project, PET was estimated using the FAO-56 Penman Monteith equation and for the historical values shown, this research employed a modified form of the Penman-Monteith equation (see Section 3.4.2.1). The FF projections typically show greater variation for each month during the baseline period and exhibit lower values in spring and higher values from August through to November compared to the historical record.



**Figure 5-8: Box plots of monthly PET totals across the Thames basin (from the historical record (based on the modified Penman-Monteith method and Met Office gridded historical weather data – see Section 3.4.2) as the wider black box-plots and the eleven-member FF ensemble over the 1961 to 2008 time period (with each individual scenario shown as a light grey box plot).**

In Figure 3-7, the historical precipitation record in terms of a rolling 365-day surplus/deficit from the long-term average 365-day precipitation (1961 to 2008) was plotted to highlight the deficits associated with sustained periods of low flow in the historical record for the Thames. Figure 5-9 presents the FF precipitation time-series (1961 to 2008) in terms of rolling 365-day surpluses/deficits from the long-term average 365-day historical precipitation (1961 to 2008) to compare the relative severity of deficits in the equivalent FF baseline periods to the historical record. This plot highlights that none of the FF precipitation sequences includes a deficit of the magnitude of that associated with the historical 1976 event, although several events are relatively close.



**Figure 5-9: Rolling 365-day deficit/surplus precipitation across the Thames basin from the eleven FF ensemble members relative to the historical LTA 365-day precipitation (E-Obs v 3.0).**

Table 5-1 presents summary information for precipitation from each FF scenario compared to the historical record over the 1961 to 2008 period, including the mean annual precipitation and the number of ‘notably’ dry periods. As for the analysis presented in Section 3.4.2.2, ‘notably’ dry periods are defined as periods when the rolling 365-day precipitation deficit is greater than 100mm (equivalent to a reduction of 15% of the 1961 to 2008 average).

These comparisons demonstrate that the FF ensemble members are generally wetter than the historical record, both in terms of the mean annual precipitation and the mean maximum deficits recorded during the notably dry periods. The exception is scenario ‘*afixl*’ which has one of the lowest estimates of mean annual precipitation and contains a dry event that is relatively close to that recorded during the 1976 period in the historical record. The frequency of dry periods is generally consistent across the FF ensemble as a whole, albeit that ‘*afixk*’ has only nine such events.

**Table 5-1: Comparison between the historical precipitation record and each FF ensemble member across for the 1961 to 2008 period for significant precipitation deficit periods. Mean and co-efficient of variance (CoV) across the FF ensemble member values are also shown. Notably dry events are these including a deficit of greater than 100mm from LTA.**

	Mean annual precipitation (mm)	Number of 'notably' dry events	Mean maximum deficit of 'notable' events (mm)	Maximum deficit (mm)
Historical record	674	14	188	325
afgcx	719	11	148	205
afixa	695	15	185	292
afixc	717	15	138	217
afixh	712	10	136	195
afixi	698	16	169	270
afixj	705	15	168	240
afixk	715	9	156	262
afixl	696	11	207	304
afixm	705	13	163	265
afixo	698	16	138	188
afixq	705	13	165	251
Mean (Future Flows ensemble)	706	13	161	244
CoV (Future Flows ensemble, %)	1	18	13	15

Table 5-2 presents the mean annual PET for each FF ensemble member and the historical record during the 1961 to 2008 period and highlights the significant variation between ensemble members. Some scenarios can be seen, on average, to be both significantly wetter (Table 5-1) and exhibit significantly lower evapotranspiration rates (Table 5-2) than the historical record (e.g. member '*afixh*') which could be expected to have implications for the general catchment water balance and the estimation of resource availability.

**Table 5-2: Comparison between mean historical PET and the mean from each FF ensemble member for the 1961 to 2008 period. Mean and co-efficient of variance (CoV) across the FF ensemble member mean values are also shown.**

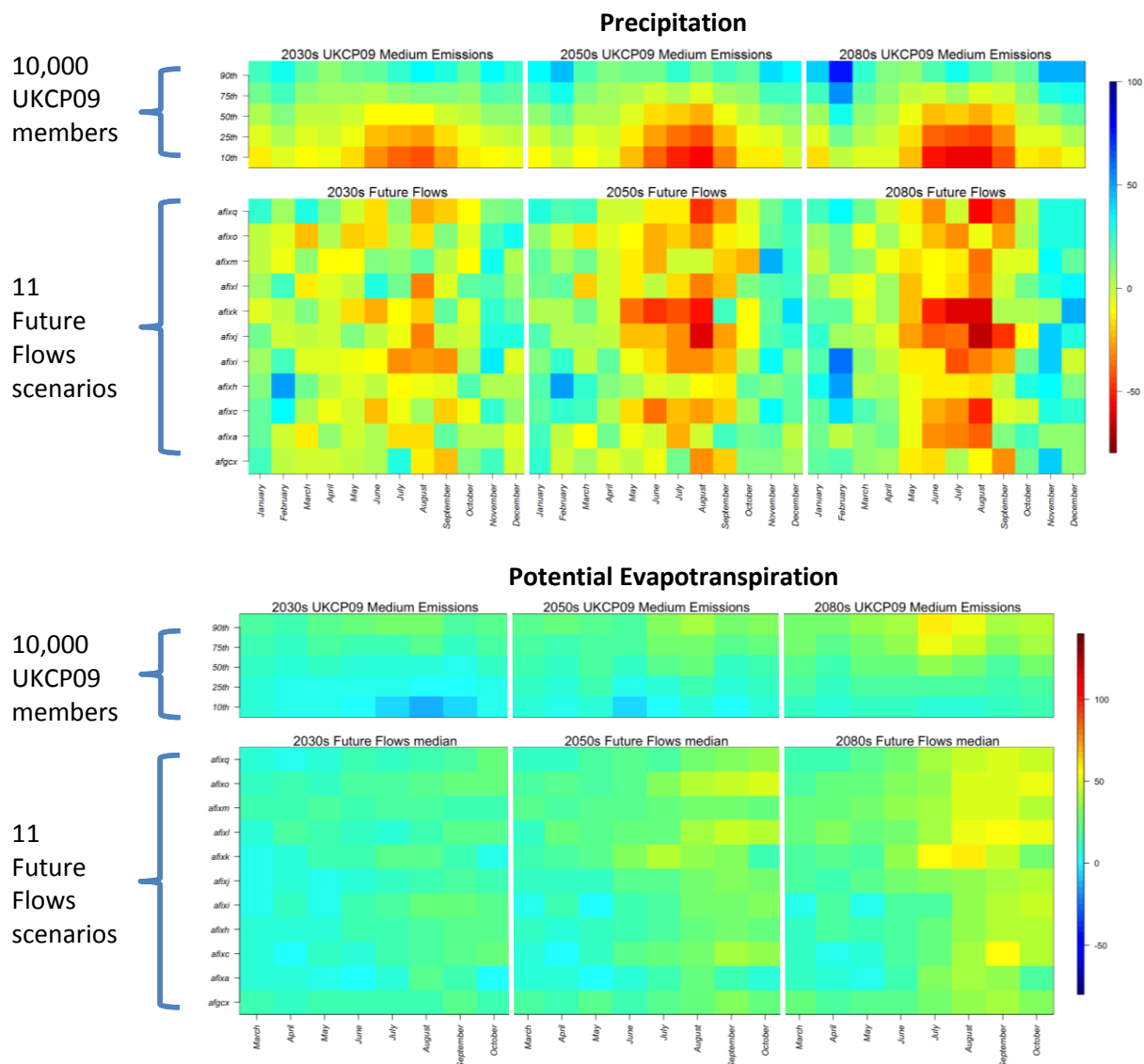
Historical (Penman)	Future Flows Ensemble Member											Ensemble	
	afgcx	afixa	afixc	afixh	afixi	afixj	afixk	afixl	afixm	afixo	afixq	mean	CoV (%)
636	609	738	688	553	636	785	700	611	643	599	661	657	10

In Section 5.4.1, the flow sequences generated using the FF 'baseline' climatological data and the historical record are compared to evaluate the information content, in terms of periods of low flows, of interest to water resources planers.

### 5.3.3 *Comparison of changes in precipitation and PET from the Future Flows and UKCP09 projections*

Figure 5-10 presents changes in 30-year time windows (1961 to 1990 to be consistent with the UKCP09 projections) from each FF ensemble member (which were driven by the Medium Emission scenario) along with the range of changes calculated from the UKCP09 Medium Emission scenarios samples. For PET, for which the modified Penman-Monteith method has been used in the plots below, only March to October months are shown due to the very large percentage changes relative to the relatively small baseline values for the other months.

Both sources of climate projections suggest drier, hotter summers and wetter winters along with a trend towards drier, hotter autumns potentially delaying the replenishment of soil moisture stores and the recharge of groundwater storage. The FF projections suggest mean precipitation reductions ranging from a reduction of 63% to an increase of 47% for specific summer months by the 2080s with the much larger UKCP09 ensemble suggesting mean changes ranging from a reduction of 65% to an increase of 37% in summer precipitation (averaged over the June – August period) noting that the eleven-member FF projections and UKCP09 projections are underpinned by the same climate models. These large ranges highlight the very high uncertainty associated with the climate projections currently available to water resources planners.



**Figure 5-10: Monthly change (as a percentage) in mean monthly precipitation (top plot) and PET (bottom plot, using the modified Penman-Monteith method) from the UKCP09 2030, 2050 and 2080 Medium Emission scenarios - 10<sup>th</sup>, 25<sup>th</sup>, 50<sup>th</sup>, 75<sup>th</sup> and 90<sup>th</sup> percentiles and each scenario from the eleven member FF ensemble. Changes are relative to a 161 to 1990 baseline.**

5.4 Hydrological modelling using the UKCP09 and Future Flows climate projections  
Section 4.3 describes the development of the multi-site model ensemble (MSME) for generating spatially coherent river flows at multiple sites of interest for water resources planning. This MSME is used in the remainder of this chapter to quantify the impacts of climate change on river flows for the study locations presented in Figure 3-1.

It is recognised that the application of the MSME ensemble for assessing the impacts of climate change assumes that the models have the same predictive skill under a changed climate as they do during the model conditioning period. Crooks & Kay (2015), in their 120-year simulation of historical rivers flows in the Thames, did demonstrate the long-term stability of parameter values (for their CLASSIC model) and suggested that there is evidence of a general stationarity of the Thames catchment hydrological response. However, the assumption of parameter stability should be noted, particularly when interpreting the MSME's use in prediction for events which may be significantly more extreme than that encountered during its development.

Prior to assessing the impacts of climate change under both the UKCP09 and FF projections, the next section compares the flow sequences for the FF 'baseline' period and the historical record in terms of the flow metrics that are of interest to water resources planners. It should be noted that for the hydrological model simulations that consider the transient FF projections, only the MSME members derived using the modified Penman-Monteith method for estimating PET (known as MSME-PM) were considered so as to be consistent with the method used as part of the development of the FF climate projections (Prudhomme et al., 2012).

#### 5.4.1 *Comparison between estimated river flow metrics using the Future Flows projections and the observed climatological record over the 1961 to 2008 period*

This section compares the Flow Duration Curves (FDCs) and Mean Summer Flows (MSFs), for low flow years, for the baseline period identified in Chapter 3 (1961 to 2008), for the historical record and the equivalent period from each FF ensemble member.

#### 5.4.1.1 Comparison of the Flow Duration Curves (FDCs) for each Future Flows (FF) ensemble member with the historical record

Figure 5-11 presents the FDC for the Thames at Kingston for MSME-PM simulations driven using the historical (1961 to 2008) climatology and the equivalent climatology from each member of the FF projections for the same time period. For clarity, only the 50<sup>th</sup> percentile estimate from each FF ensemble member is shown. This FDC highlights that for higher flows, the MSME-PM 50<sup>th</sup> percentile estimate of flows from all members of the FF ensemble are typically located in the lower half of the distribution of those generated using the historical climatological record. During the medium to low flow range (Q30 to Q99), the FF ensemble members are located across the full range of values from the historical record with a couple of FF ensemble members notably higher.

To quantitatively assess the similarity between the conditional probabilities generated from the MSME-PM driven by the historical climatology with each FF ensemble member, a simple measure, based on the Combined Overlap Percentage (COP) measure used by Westerberg et al. (2011a), is presented in equation 5.1. This measure calculates the mean of the percentage of the overlapping range of the FDC conditional uncertainty bands driven by the observed climatology with each FF ensemble member (Equation 5.1). This measure is designed to ensure that a perfect match of 100% for COP cannot be achieved if the simulated uncertainty is overestimated.

$$\frac{\sum_{t=1}^T \left( \text{mean} \left( \frac{QR_{\text{overlap}}}{QR_{\text{obs}}}, \frac{QR_{\text{overlap}}}{QR_{\text{sim}}} \right) \right)}{T} \quad (5.1)$$

where  $T$  is the number of points along the FDC considered,  $QR_{\text{overlap}}$  the intersection between the 'observed' and the FF FDC ranges,  $QR_{\text{obs}}$  the observed FDC range and  $QR_{\text{FF}}$  the FF FDC range.

This measure is presented for each FF ensemble member in Table 5-3. The first value is the score calculated for nine points along the entire FDC (Q1, Q5, Q10, Q30, Q50, Q70, Q90, Q95 and Q99) as  $COP_{FDC\_Full}$ . The second score only considers the lower half of the flow duration curve (Q50, Q70, Q90, Q95 and Q99), as  $COP_{FDC\_Low}$ , to focus on the part of the FDC of greater interest to the water resource planner. A score of 100 percent means that for each point along the FDC that the conditional uncertainty ranges exactly match. A score of 0 percent would indicate that there is no overlap at any of the FDC points considered.

**Table 5-3: Combined Overlap Percentage scores for each FF ensemble member compared to that using the historical climatology. Score based on selected flow percentiles (Q1, Q5, Q10, Q30, Q50, Q70, Q90, Q95 and Q99) over the 1961 to 2008 period.**

Measure	Future Flows Ensemble Member										
	afgcx	afixa	afixc	afixh	afixi	afixj	afixk	afixl	afixm	afixo	afixq
$COP_{FDC\_Full}$	55	59	63	30	73	69	65	68	78	63	69
$COP_{FDC\_Low}$	44	63	61	7	77	61	69	64	89	61	70

These results highlight that the MSME-PM simulated FDCs for some of the FF projections are quite different from that generated using the observed climatology. FF ensemble member ‘*afixh*’ in particular is very different to the observed FDC, and this divergence with the observed FDC is a reflection that it is a relatively wet scenario with low mean rates of PET compared to the historical record (see Table 5-1 and Table 5-2). In contrast, member ‘*afixm*’ scores highly under both measures, especially so at the lower flows, and it can be seen from Table 5-1 and Table 5-2 that it is slightly wetter than the observed record but also has slightly higher mean rates of PET to compensate for this increased wetness.





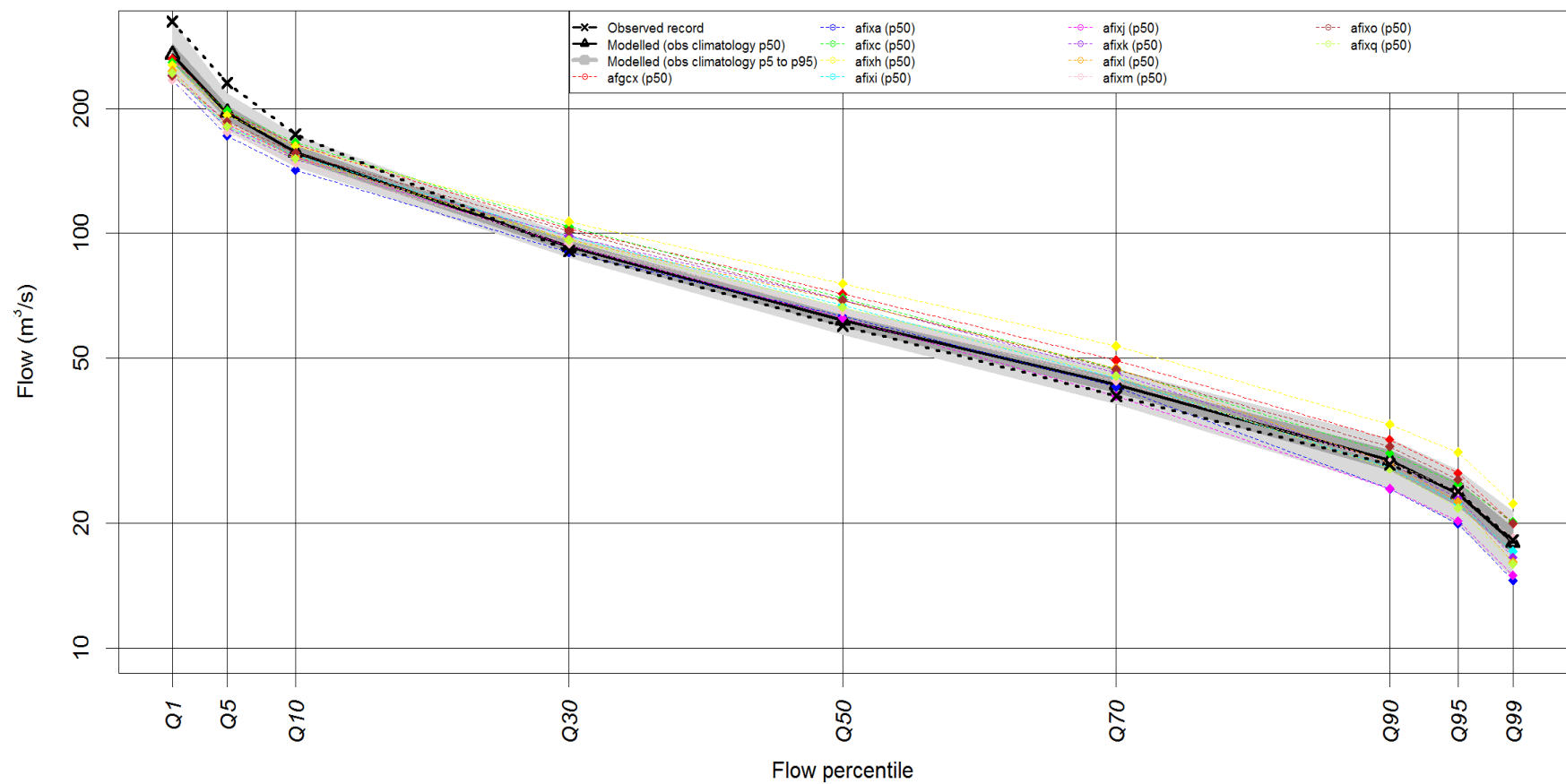


Figure 5-11: MSME-PM simulated FDC for Kingston for the 1961 to 2008 climatology from each FF ensemble member (50<sup>th</sup> percentile for each FF scenario shown as a distinct colour) and the historical 1961 to 2008 climatology (shown as grey shared area [5<sup>th</sup> to 95<sup>th</sup> percentile] and black triangles [50<sup>th</sup> percentile]). Values from the historical flow record (1961 to 2008) shown as black crosses.

#### 5.4.1.2 Assessment of the reproduction of low flow sequences

Thames Water's Drought Plan considers the average naturalised flow in the River Thames (at Kingston) between April to September to provide a good approximation to the water resources available in the London WRZ (Thames Water, 2010).

Figure 5-12 presents the historical lowest Mean Summer Flow (MSF, April to September) at Kingston (39001) for the five lowest years in the 1961 to 2008 along with the conditional probabilities of flows using the MSME-PM driven by the historical climatology. The 50<sup>th</sup> percentile MSME-PM predictions for the five years with the lowest MSF from each FF ensemble member during the 1961 to 2008 period are also shown as a comparison. The plot highlights that none of the FF ensemble members contain periods of MSF that are as severe as the historical record. There are some FF projections which do not contain any flow sequences that are as severe as the 5<sup>th</sup> most severe event generated using the historical climate record (e.g. 'afixh', 'afixc' 'afixo') and there is a tendency for some members (e.g. scenario 'afixa') to contain relatively lower low flow periods compared to other FF ensemble members. This may suggest that the natural climate variability and the presence of persistent dry periods within the FF ensemble may be limited which is of particular concern in terms of their application for water resources planning.

#### 5.4.2 Assessing the impacts of climate change on river flows for water resources planning

This section explores the projected changes in selected flow percentiles and changes to MSFs for low flow years using both the UKCP09 and FF projections. The UKCP09 projections are relative to a 1961 to 1990 baseline and therefore the projected changes under the FF projections use the same reference baseline period (rather than the 1961 to 2008 period considered previously).

The estimated impacts of climate change under both the FF and UKCP09 projections are considered in this section, with the impacts under the UKCP09 projections based on the 20-member sub-samples presented earlier in this chapter using both methods for estimating PET (MSME simulations). Each FF ensemble member scenario is presented individually alongside the range of impacts under UKCP09 – noting that for the FF projections, only the hydrological model ensemble members derived using the modified Penman-Monteith method for estimating PET (MSME-PM simulations) have been used.

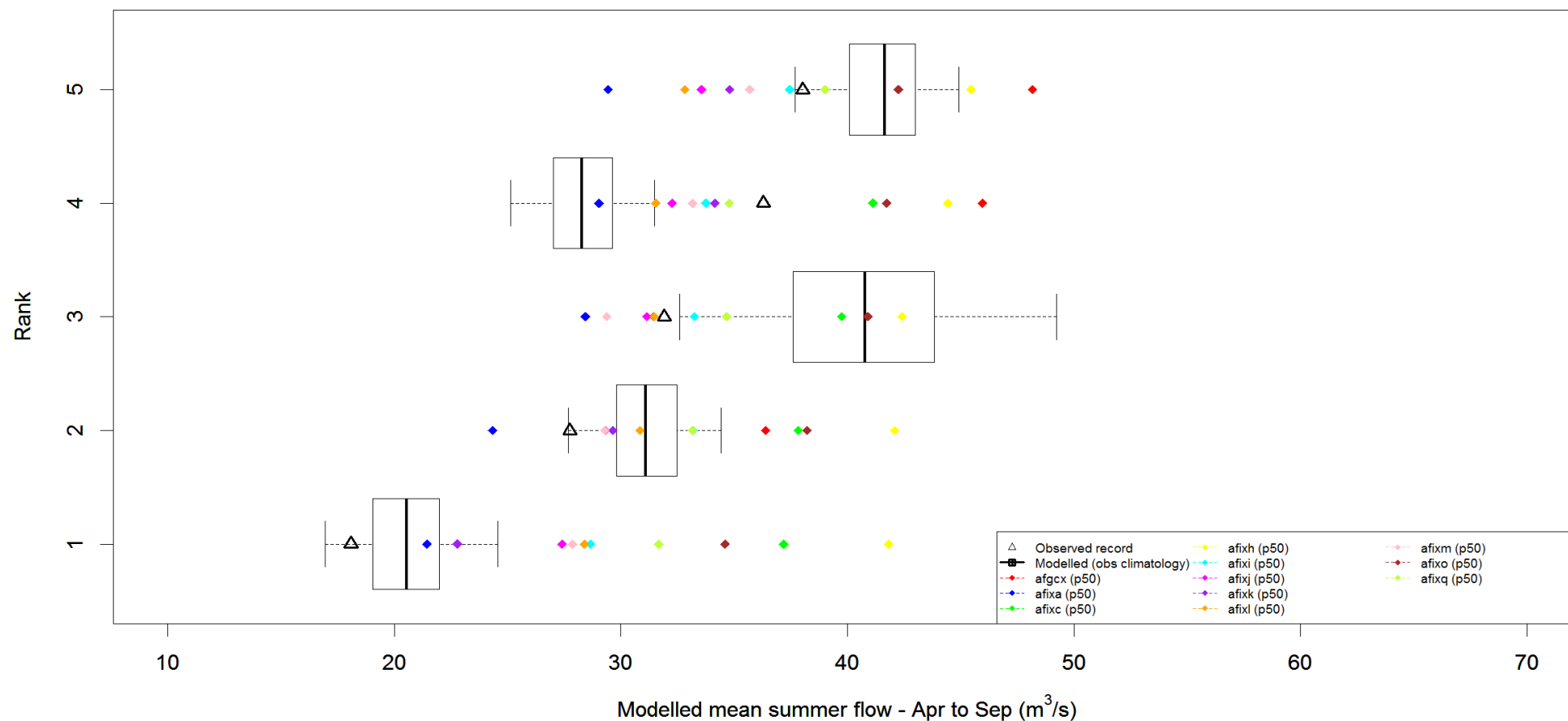


Figure 5-12: Lowest MSF at Kingston for the lowest years in the 1961 to 2008 historical record (black triangles). Distribution of simulated values (black box-plots) and the 50<sup>th</sup> percentile MSME-PM simulated value for the five lowest MSF values from each FF ensemble member (each FF scenario shown as a distinct colour) also shown.

#### 5.4.2.1 Impacts on selected flow percentiles from the Flow Duration Curve

Figure 5-13 presents the percentage change, relative to a 1961 to 1990 baseline, in Q95, Q70 and Q50 for the Thames at Kingston under the UKCP09 sub-samples (MSME simulations) and each FF ensemble member (MSME-PM simulations) for the 2030s, 2050s and 2080s under a Medium Emission scenario. In terms of a trajectory of impacts over time, the UKCP09 projections suggest increasing reductions in low to medium flows through to the 2080s. As an example, the central estimates for the changes in Q95 are reductions of 13.2%, 14.1% and 21.3% for the 2030s (equivalent to 0.22% reduction per year over 60 years from the 1970s), 2050s (equivalent to 0.45% reduction per year over 20 years from the 2030s) and 2080s (equivalent to 0.24% reduction per year over 30 years from the 2050s) respectively. Table 5-4 presents a summary of the results for the 2080s.

**Table 5-4: 5<sup>th</sup>, 50<sup>th</sup> and 95<sup>th</sup> predicted change (percentage relative to a 1961 to 1990 baseline) using the MSME-PM (Future Flows) and MSME (UKCP09) in Q95, Q70 and Q50 for the Thames at Kingston under the UKCP09 sub-sample and each FF ensemble member for the 2080s.**

Flow Percentile	Future Flows (MSME-PM)		95 <sup>th</sup> Percentile	UKCP09 (MSME)	
	Lower Bound	Upper Bound		50 <sup>th</sup> Percentile	5 <sup>th</sup> Percentile
Q95	-3.6%	-28.1%	-1.4%	-21.3%	-35.3%
Q70	-0.0%	-28.3%	-1.2%	-23.3%	-36.9%
Q50	+16.2% (' <i>afixh</i> ')	-22.4%	+0.8%	-15.6%	-29.8%

For the FF ensemble, these results demonstrate the wide variation across the different FF ensemble members and that the relative ranking of impacts across the eleven members can differ significantly depending upon the time-horizon being considered which is likely to be due to the relative severity and frequency of low flow periods in the respective time-windows. The individual FF ensemble members generally sit within the 5<sup>th</sup> to 95<sup>th</sup> percentile ranges of the UKCP09 projections. The exception is FF ensemble member '*afixh*' which, at Q50, projects significant increases in flows relative to the range of projected impacts under the UKCP09 projections.

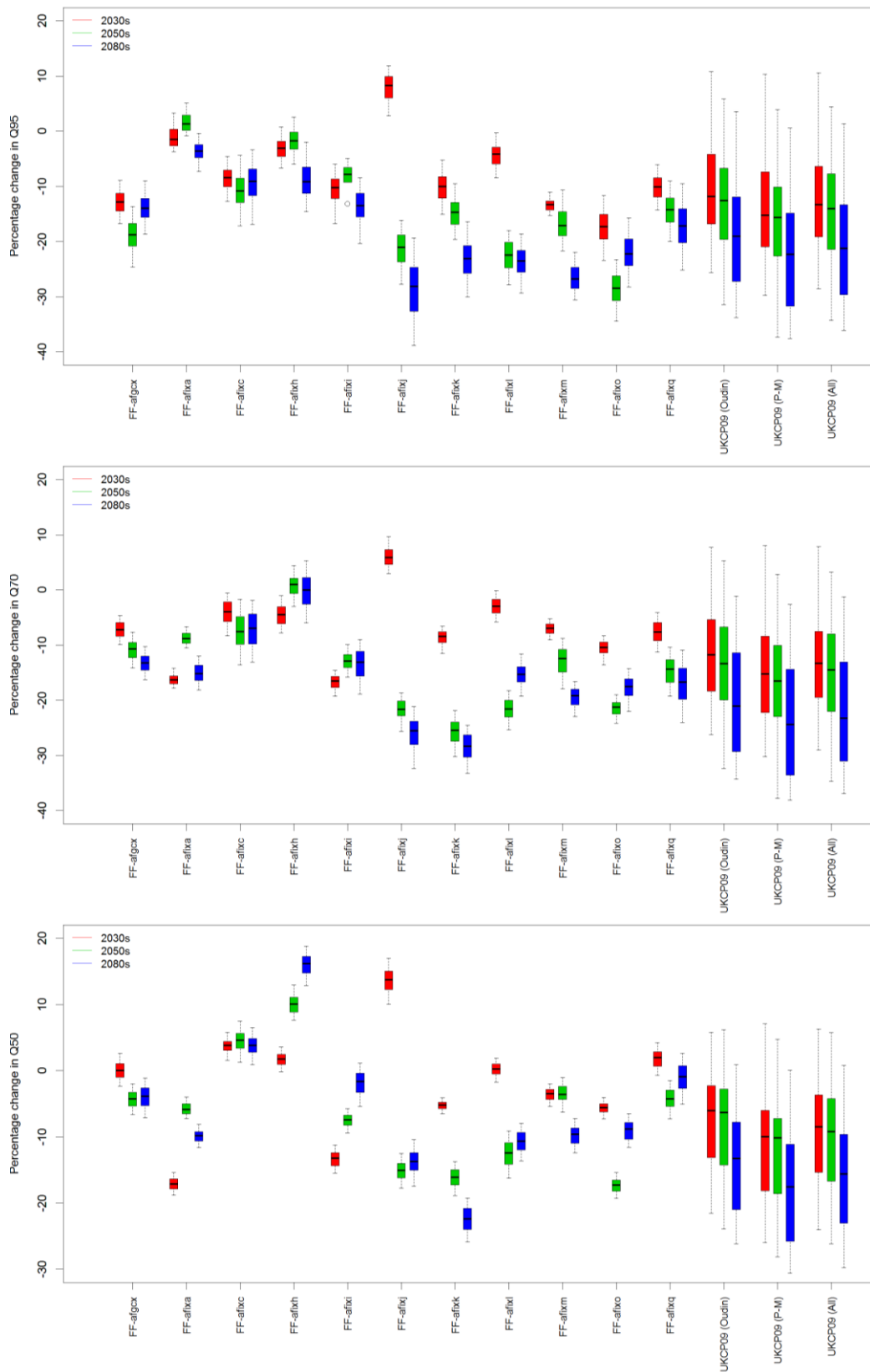
This is consistent with the changes presented in the FF reporting (Prudhomme et al., 2012) in which, for the Thames, this FF ensemble member ('*afixh*') suggested significantly higher flows than the other FF ensemble members. This is also consistent with the evidence presented earlier in this chapter that, particularly for the 2080s, this FF scenario is relatively wet and cool (see Figure 5-10). In addition, this FF scenario is the wettest during the baseline period with relatively high values for precipitation (Table 5-1) and low values for PET (Table 5-2) resulting in this FF ensemble member generating the highest flows (Figure 5-11). This FF ensemble member also scored lowest for the Combined Overlap Percentage measure for the FDC (Table 5-3). All this

evidence suggests that confidence (i.e. in terms of its ‘fitness-for-purpose’) in this FF ensemble member may be considered relatively low.

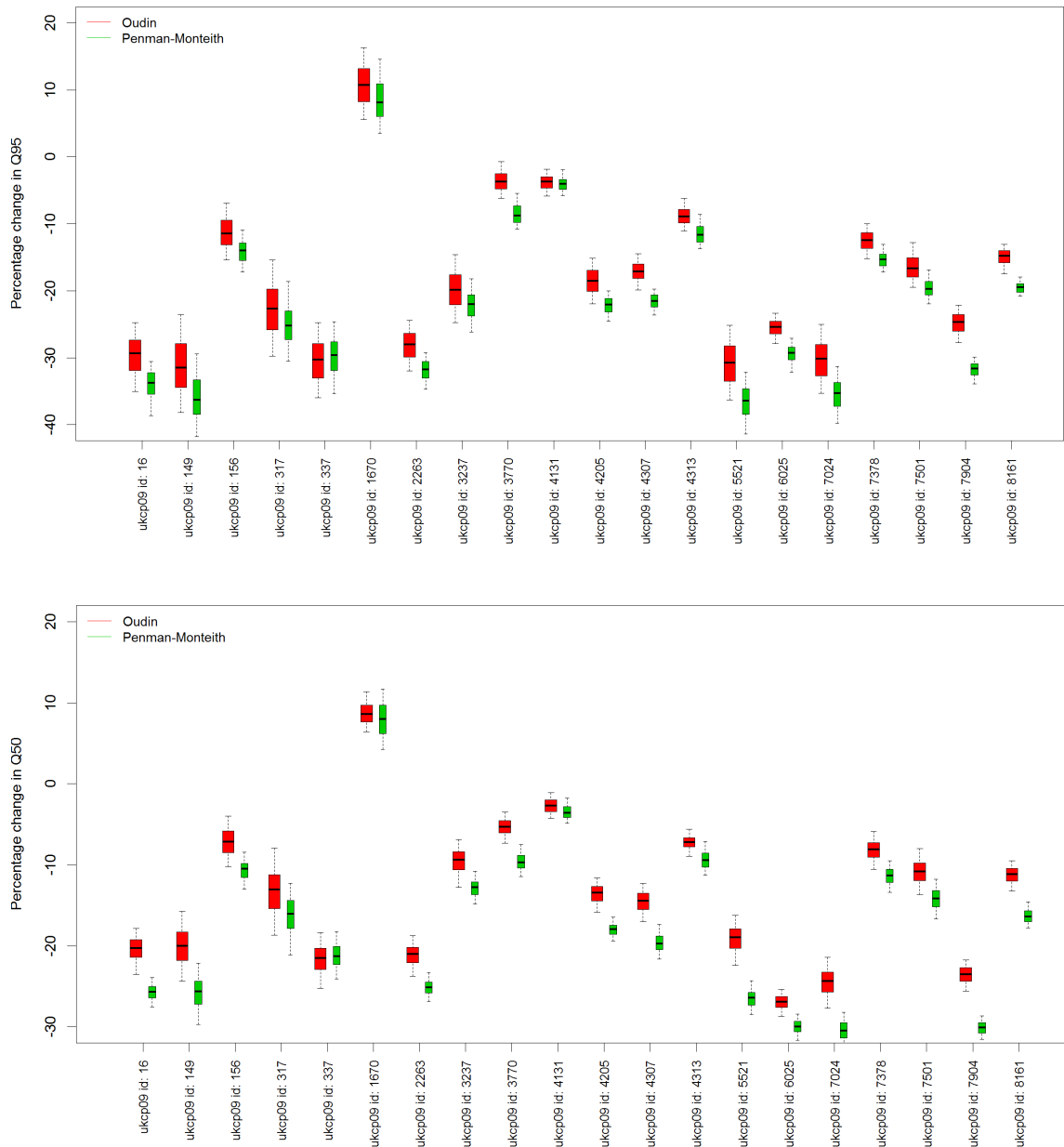
When looking at the differences due to the method used for estimating PET, it can be seen that simulations using the modified Penman-Monteith method (MSME-PM simulations) consistently predict larger reductions in flows. At Q95, for the 2080s, the range and central estimate (5th, 50th and 95th percentiles) of projected percentage changes using the Oudin method (MSME-OU simulations) are +3.6%, -19.1% and -33.8% respectively and +0.6%, -22.3% and -37.6% respectively when using the modified Penman-Monteith method (MSME-PE simulations).

Figure 5-14 further explores the difference between the two different methods used to estimate PET for the UKCP09 projections. This figure presents the percentage change in Q95 (Figure 5-14a) and Q50 (Figure 5-14b) for each sample member from the UKCP09 sub-sample for the 2080s. This figure highlights that the difference due to the PET method varies from sub-sample member to sub-sample member. For example, the central estimate of potential change under UKCP09 scenario ‘7904’ (for the 2080s) is a reduction of 24.7% using the MSME-OU and a reduction of 31.6% using the MSME-PM. The relative difference in impacts reflects the contribution from changes in precipitation and PET on flows with the largest difference under those scenarios with the greatest predicted increase in summer PET (see Figure 5-5).

The plots also highlight the relative uncertainties associated with the climate change projections relative to the uncertainty associated with the hydrological model parameters, with the results suggesting that for all time-horizons considered, the former is much larger (as shown by the variation between climate scenarios) but that hydrological parameter uncertainty is still significant (i.e. reflected in the variation within each individual climate scenario). This is consistent with the evidence reported from other similar studies (e.g. Wilby and Harris, 2006; New et al., 2007a; Manning et al., 2009; Christerson et al., 2012; Charlton & Arnell, 2014).



**Figure 5-13: The range of predicted changes at top) Q95, middle) Q70 and bottom) Q50 at under each FF scenario (MSME-PM simulations) and the UKCP09 20-member sub-samples (MSME simulations) for the 2030s (red), 2050s (green) and 2080s (blue) relative to a 1961-1990 baseline.**



**Figure 5-14: Change (as a percentage) in top) Q95 and bottom) Q50 at Kingston for individual UKCP09 scenario members for the 2080s (relative to a 1961-1990 baseline) using the MSME-OU (red) and MSME-PM (green).**

As reported by Charlton & Arnell (2014), there is little published evidence exploring the potential impacts of climate change on medium to low flows in the Thames catchment under the UKCP09 projections, particularly beyond the 2020s (see Christerson et al., 2012), with which to compare the above findings. The literature review undertaken as part of this research has also highlighted a similar lack of evidence for the application of the FF projections outside of the FF project itself (Prudhomme et al., 2012), although some studies (e.g. Cloke et al., 2010; Cloke et al., 2013) have used the underlying RCM outputs (see Section 2.3.2).

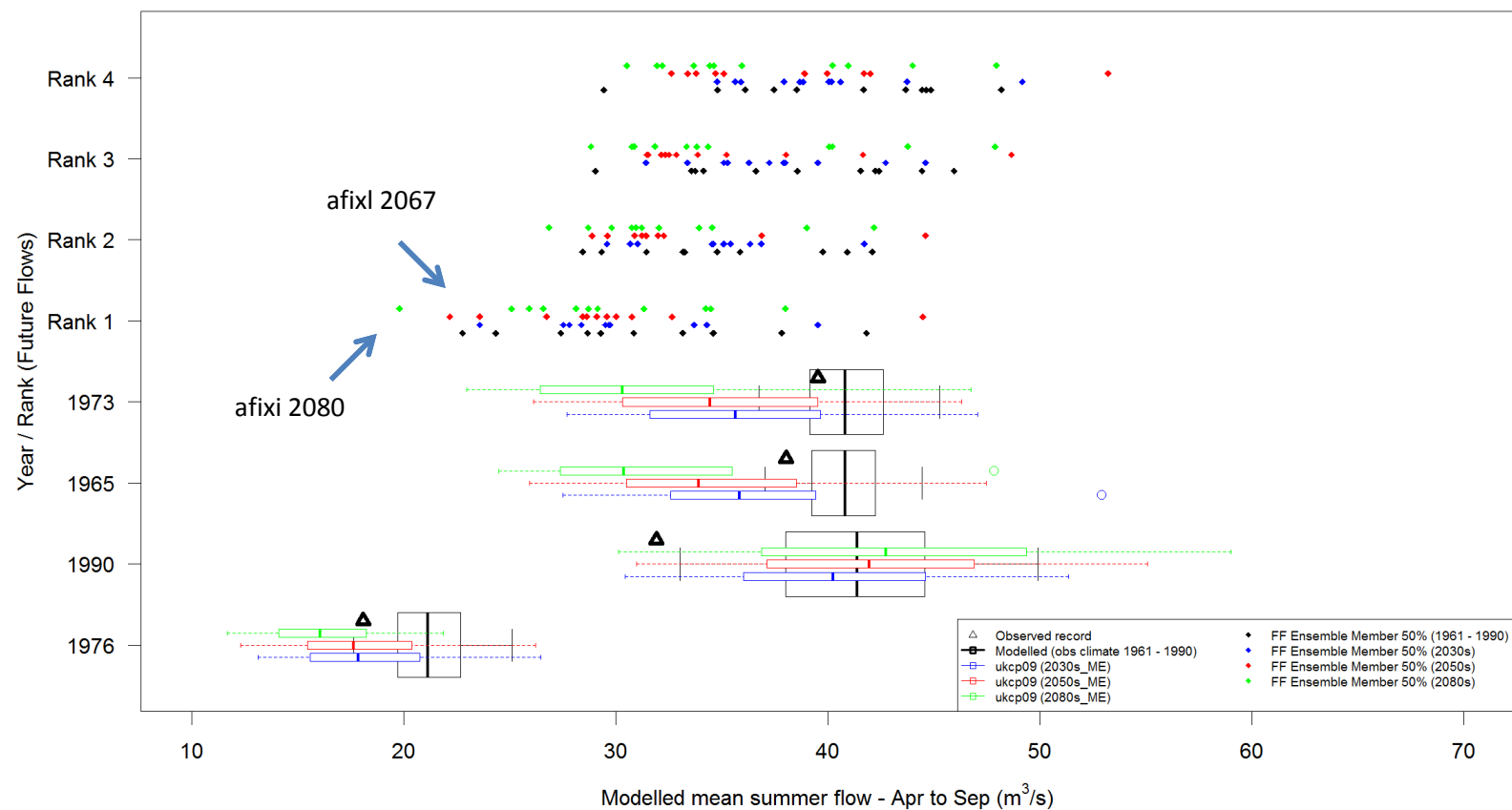


Two studies which have also made use of the UKCP09 derived products to look at climate change impacts on a permeable basin in the south-east of England are the research by Cloke et al. (2010) and Charlton & Arnell (2014). On the permeable Medway catchment, using CatchMod and a 10-member bias-correction ensemble of RCM outputs from the UKCP09 projections, Cloke et al. (2010) suggested a change in Q95 of between -15 and -35% over the 2020 to 2050 period compared to a 1960 to 1990 baseline. For the same catchment, Charlton & Arnell (2014), using the CAT-PDM hydrological model, suggested changes between around -20 and -60% reductions in Q95 by the 2080s (noting that these have been taken by 'reading off' these values from the plots provided in the paper). These impacts are larger than reported here but it should be noted that these are for a different catchment, using different hydrological model ensembles and different approaches to quantifying hydrological modelling and climate change uncertainties and so direct comparisons are difficult.

#### 5.4.2.2 Impacts on Mean Summer Flows (MSF)

Figure 5-15 presents the modelled lowest MSF (April to September) at Kingston (39001) for the four lowest years in the 1961 to 1990 (not for the longer 1961 to 2008 period considered in Figure 5-12) historical record along with the conditional probabilities of estimated flows using the MSME (UKCP09 projections) and MSME (FF projections) driven by the historical climatology. In addition, this plot presents the range and distribution of MSME estimates from the UKCP09 sub-samples for the 2030s, 2050s and 2080s all under a Medium Emission scenario with each sub-sample member considered equally likely. Finally the MSME-PM central estimate (50<sup>th</sup> percentile) estimated flow from each FF ensemble member for the four years with the lowest MSF is also shown. Two specific periods of low flow years are also highlighted (scenario afixl in 2067 and scenario afixi in 2080).

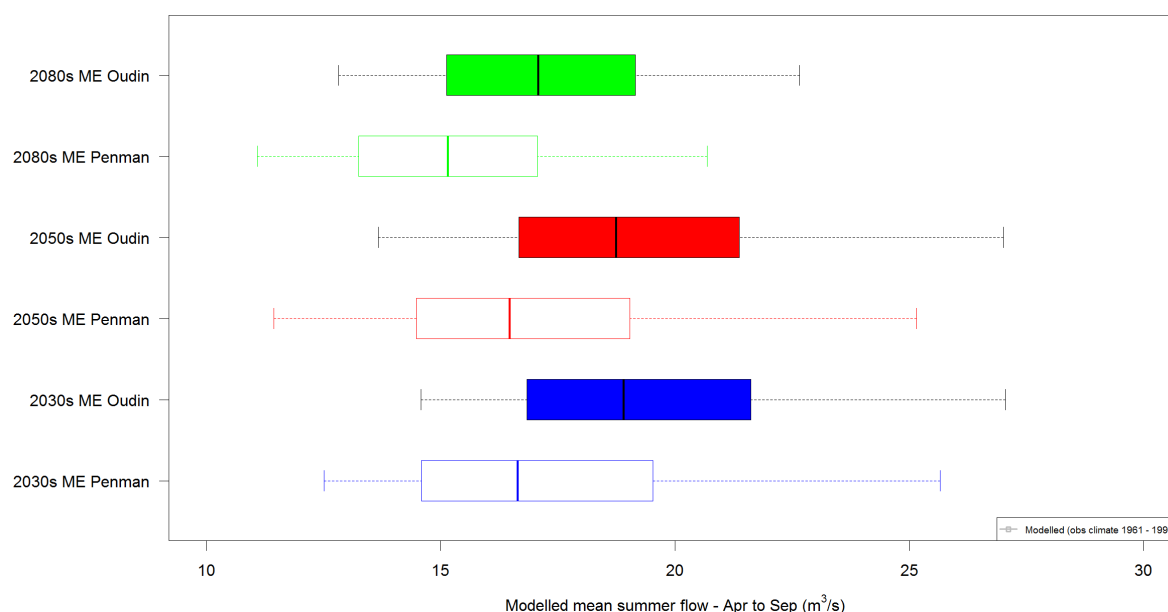
The MSME simulated results for the UKCP09 scenarios demonstrate a significant reduction in the MSF during the majority of these events, particularly for the 2080s-perturbed 1976 sequence. Using the baseline climatology, the modelled mean summer flow ranges from 17.6m<sup>3</sup>/s (5<sup>th</sup> percentile) to 25.1m<sup>3</sup>/s (95<sup>th</sup> percentile) with a central estimate of 21.1m<sup>3</sup>/s [noting that the value from the historical flow record is 18.1m<sup>3</sup>/s]. By the 2080s, the modelled MSFs, under the UKCP09 sampled scenarios, range from 11.7m<sup>3</sup>/s (5<sup>th</sup> percentile) to 21.9m<sup>3</sup>/s (95<sup>th</sup> percentile), with a central estimate of 16.1%.



**Figure 5-15:** Historical lowest MSF at Kingston for the four lowest years in the 1961 to 1990 historical record (black triangles) along with the distribution of simulated values driven by the historical climatology (black box-plots). The equivalent sequences from the UKCP09 sub-sample (MSME simulations, shown as black box-plots) and FF (MSME-PM simulations, 50<sup>th</sup> percentile only shown for each FF scenario as solid circle) for the 2030s (blue), 2050s (red) and 2080s (green).

The perturbed 1990 sequence is interesting in that the range of future values is little altered from the baseline. However, Figure 4-10 highlighted that the MSME simulations over-estimated this sequence significantly which may be related to challenges for the MSME in reproducing the recession during the summer of 1990, when winter precipitation appeared to sustain river flows (via a baseflow contribution) into the summer to a greater extent than suggested by the historical flow record. With the UKCP09 scenarios having a tendency towards increased winter precipitation, this may be offsetting the increased PET and decreased precipitation due to climate change that takes place during the summer.

Figure 5-16 focuses on the UKCP09 projections and compares the two alternative methods for estimating PET and examines the results for the perturbed 1976 flow sequence. This plot highlights the significant difference between the two PET methods, with the 5<sup>th</sup>, 50<sup>th</sup> and 95<sup>th</sup> conditional probabilities 11.1, 15.2 and 20.7m<sup>3</sup>/s respectively under the MSME-PM simulations for the 2080s and 12.8, 17.1 and 22.7m<sup>3</sup>/s respectively using the MSME-OU for the 2080s. Similar differences due to the PET method are evident for the 2030s and 2050s time-horizons. The potential implications of this for water resources planning are considered in Chapter 6.



**Figure 5-16: Comparison between MSME-OU (solid shading) and MSME-PM (no shading) simulations for the UKCP09 20-member sub-samples in terms of the impact on the lowest MSF event (the perturbed summer of 1976) for the 2030s (blue), 2050s (red) and 2080s (green).**

These results highlight that none of the FF scenarios contain MSFs that are as severe as those resulting from the UKCP09 perturbed scenarios. Only a single event, from one FF ensemble member during the 2080s period, containing a sequence as severe (in terms of mean summer flow) as present when using the baseline historical climate record. This again suggests that the natural climate variability, and the presence of persistent dry periods, within the FF ensemble may be limited and is consistent with the concerns previously reported (e.g. Rocheta et al., 2014; Watts et al., 2015) that the outputs from the climate models that were used to develop the UKCP09 probabilistic and FF projections have limitations in reproducing the blocking patterns that are significant in the formation of extended droughts. Some FF scenarios are again shown to be particularly wet, with the lowest MSF from some members during the 2080s greater than all four of the summers shown from the historical record.

Two specific sequences that could present challenges to the London water supply system are listed in Table 5-5. These are the summer of 2080 from FF ensemble member '*afixi*' and 2067 from member '*afixl*', both also shown in Figure 5-16. The 5<sup>th</sup>, 50<sup>th</sup> and 95<sup>th</sup> percentile conditional estimates associated with these two events are also presented in Table 5-5.

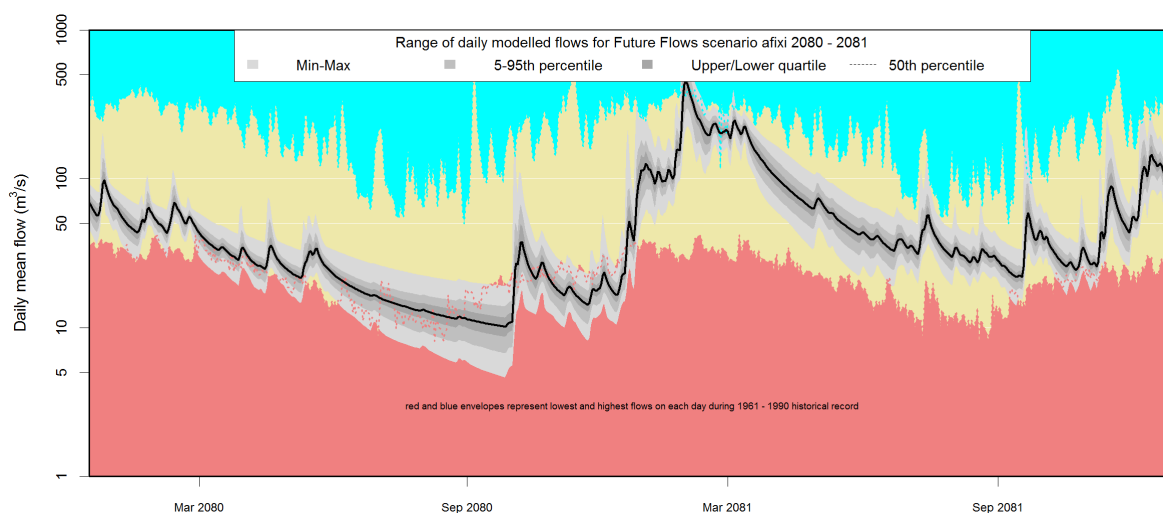
**Table 5-5: Potential events of interest for water resources planning in London in the FF ensemble.**

'Event'	Future Flows Conditional Flow Estimates of MSF (m <sup>3</sup> /s)		
	95 <sup>th</sup> Percentile	50 <sup>th</sup> Percentile	5 <sup>th</sup> Percentile
'afixi' 2080	16.6	19.8	23.2
'afixl' 2067	18.6	22.2	25.6

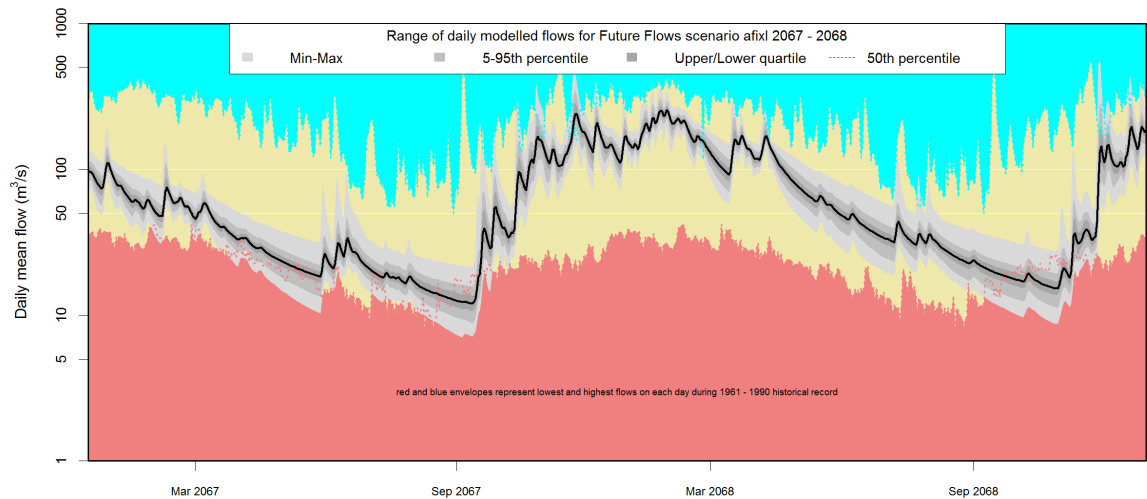
Figure 5-17 presents the hydrograph generated using the MSME-PM simulations for the 2079 to 2080 period in the FF ensemble member '*afixi*' relative to the minimum and maximum values in the historical record between 1961 and 1990. This is the most 'severe' MSF period from all the FF scenarios as measured by the MSF from April to September although the MSF is still higher than that experienced in 1976 in the historical record. However, on closer examination of the hydrograph, it can be seen that from mid-summer 2080 onwards flows are progressively lower relative to the historical record with only a modest, delayed, recovery in autumn and early winter 2081 before a full 'recovery' during late winter and early spring.

Another interesting sequence within the FF scenarios is the period 2067 to 2068 from the '*afixl*' Future Flows scenario, with the hydrographs from the MSME-PM simulations for this sequence presented in Figure 5-18. This also sequence exhibits reduced flows over this two year period. The severity, relative to the 1976 historical event, of both of the above two sequences is not necessarily reflected in the MSF metric used by Thames Water (2010).

Across both plots the range of the conditional probabilities highlight the sensitivity of modelled flows during such periods is relatively large (albeit noting the logarithmic scale which visually magnifies the differences at low flows). Another feature is that the periods of low flow associated with these droughts persist further into the autumn and winter periods than was experienced in 1976. As a result, considering only the MSF from April to September (as used by Thames Water, 2010) may not be an adequate indicator as to the relative significance of potential drought events in the future which may not follow similar patterns to those in the historical record.



**Figure 5-17: MSME-PM simulated hydrograph at Kingston for 2080 to 2081 for the 'afixi' FF scenario. The blue and red shaded areas (and dotted red line when covered by the modelled daily flow percentiles) show the extent of the historical envelope of daily maximum and minimum flows during the historical flow record 1961 to 1990.**

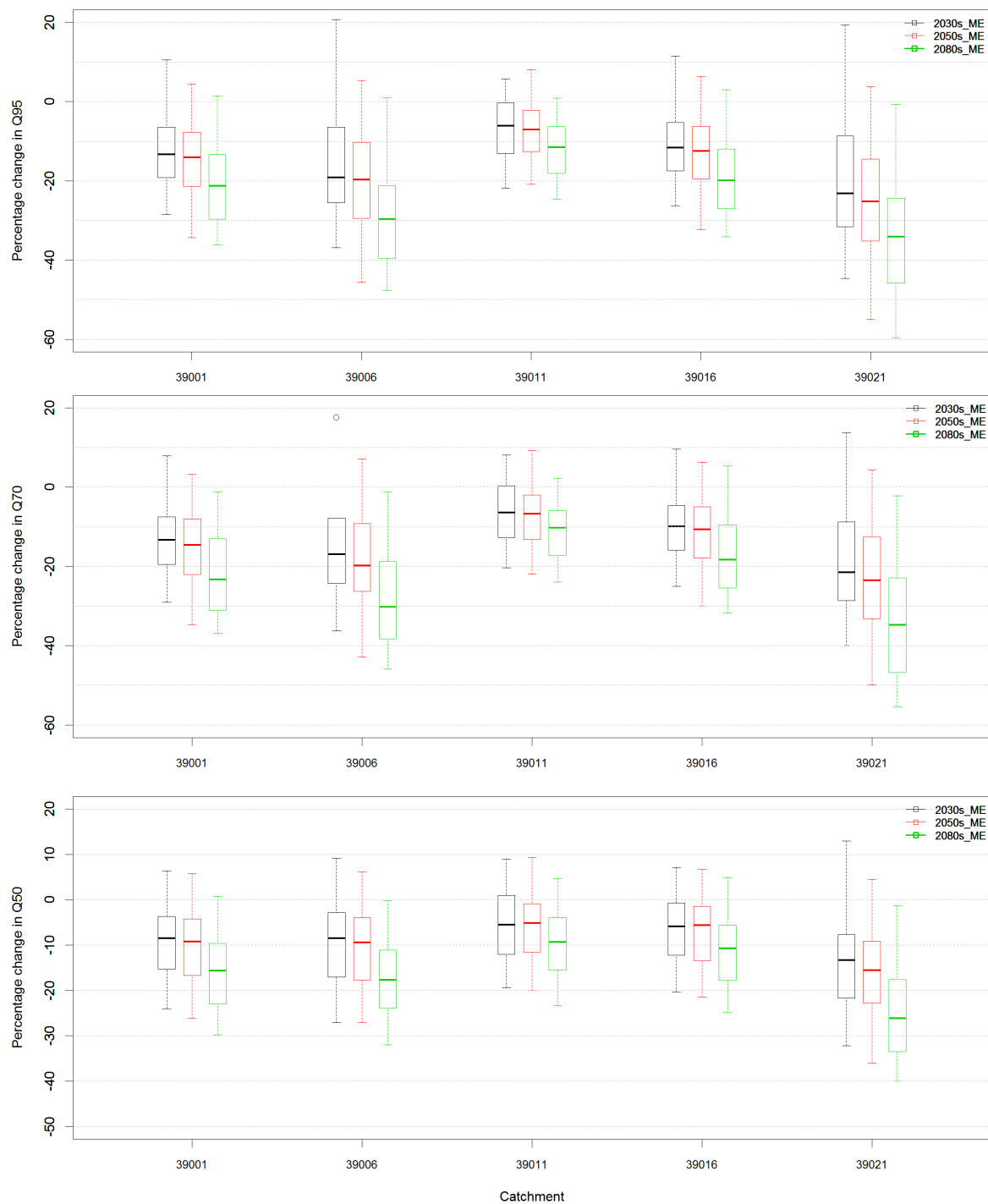


**Figure 5-18: MSME-PM simulated hydrograph at Kingston for 2067 to 2068 for the ‘afixl’ FF scenario. The blue and red shaded areas (and dotted red line when covered by the modelled daily flow percentiles) show the extent of the historical envelope of daily maximum and minimum flows during the historical flow record 1961 to 1990.**

#### 5.4.2.3 Impacts across sub-catchments in the Thames

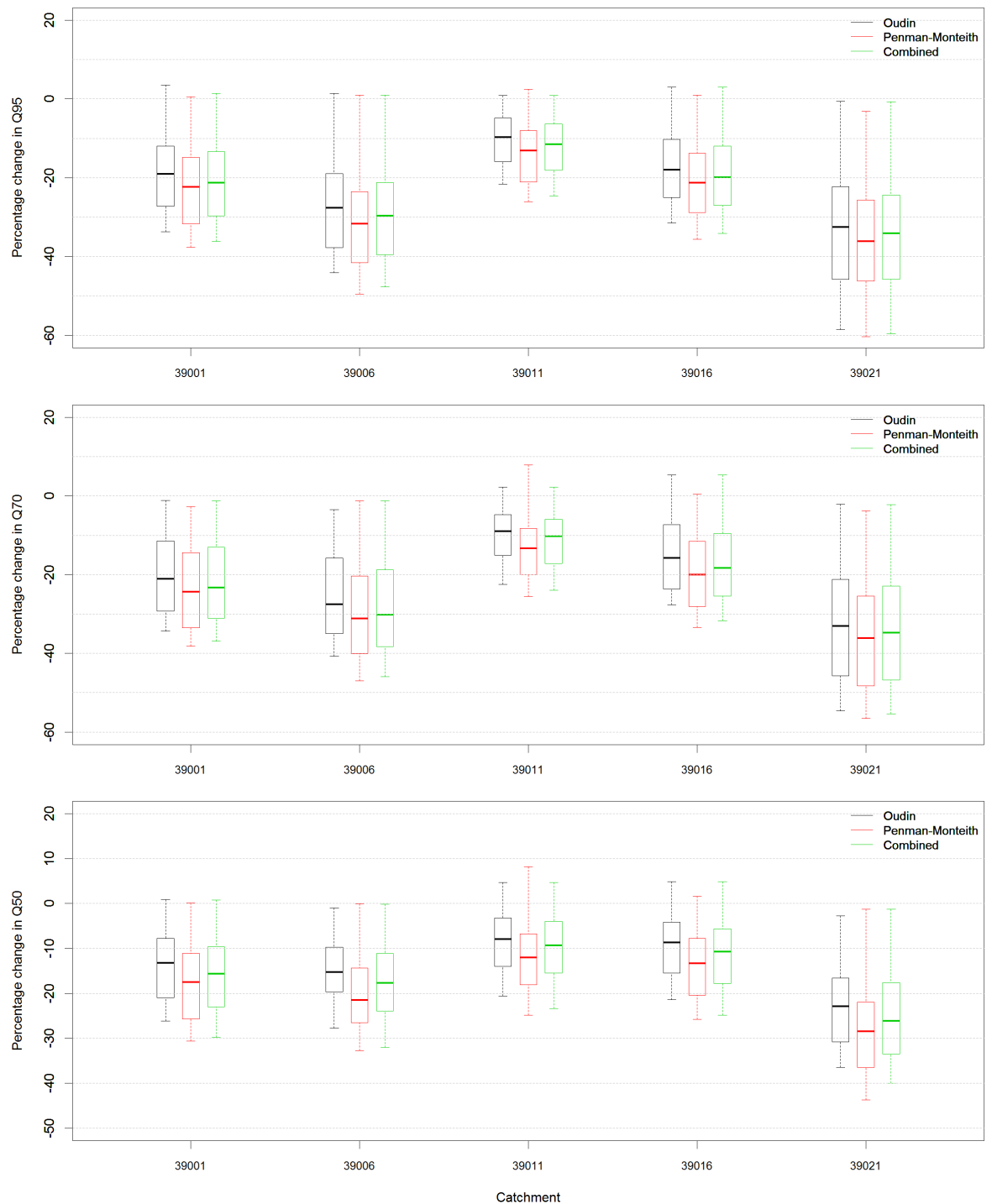
Figure 5-19 presents the MSME simulated changes to Q95, Q70 and Q50 across the five study catchments taken forward into the climate change impacts analysis. The results suggest a similar magnitude of impacts across these catchments with Enslow Mill (39021) generating the largest reductions. Enslow Mill (39021) is a flashier, smaller catchment dominated by the aquitard hydrogeological class (see Section 3.3), with a relatively low baseflow index (BFI, after Gustard et al., 1992). This suggests that smaller, flashier catchments may suffer impacts to low flows to a greater degree than baseflow supported regimes that benefit from the underground storage of winter rainfall to support flows during the summer. This is consistent with the results from other studies, such as Charlton & Arnell (2014). Charlton & Arnell (2014) highlighted that in lowland, relatively permeable catchments, low flows are largely generated by drainage from water stored in soil and groundwater. In contrast, for flashier, less permeable catchments, low flows are relatively more dependent upon direct runoff from precipitation events and therefore such catchments are likely to be more sensitive to climate change projections.

The results previously presented for the Thames catchment to Kingston (39001) suggested a significant influence from the method used to calculate PET on the projected impacts of climate change. Figure 5-20 further examines this influence across the five study catchments for the 2080s under a Medium Emission scenario and confirms a consistent difference between the two methods irrespective of the catchment under consideration.



Catchment	39001	39006	39011	39016	39021
BFI	0.63	0.86	0.72	0.88	0.65

**Figure 5-19: Comparison of the MSME simulated impacts on Q95 (top), Q70 (middle) and Q50 (bottom) for different sub-catchments using the UKCP09 climate projections for 2030s (black box-plot), 2050s (red box-plot) and 2080s (green box-plot), all relative to a 1961 to 1990 baseline under a Medium Emission scenario. Catchment BFI (Gustard et al., 1992) also presented.**



Catchment	39001	39006	39011	39016	39021
BFI	0.63	0.86	0.72	0.88	0.65

**Figure 5-20: Comparison of the simulated impacts on Q95 (top), Q70 (middle) and Q50 (bottom) for different sub-catchments using the UKCP09 climate projections for the 2080s (relative to a 1961 to 1990 baseline) under a Medium Emission scenario, with results presented for MSME-OU (black box-plot), MSME-PM (red box-plot) and MSME (green box-plot) simulations respectively. Catchment BFI (Gustard et al., 1992) also presented.**



## 5.5 Chapter summary

The research presented in this chapter has examined the two climate change products recommended for use in water resources planning in England and Wales (UKCP09 and Future Flows (FF) climate change products) in the context of water resources planning. The Multi-Site Model Ensemble described in 4.3.4 has been used to model the impacts of climate change on river flows and evaluate the relative significance of different sources of uncertainty (in this thesis namely hydrological model parameter uncertainty, climate model uncertainty and the method used to estimate PET) for water resources planning in the Thames basin.

### *Comparison of 'baseline' (1961 to 2008) climatology and river flows between the Future Flows ensemble and the historical record*

The guidance that accompanies the FF projections suggests that each member should be considered equally plausible (Prudhomme et al., 2012) but relevant evidence as to the characteristics of each member is not readily available to water resource planners. In this research, the 'baseline' period (1961 to 2008) identified in Chapter 3 from each of the FF ensemble members was compared to the historical climatological record prior to their application to project future river flows. This comparison highlighted that the range and magnitude of monthly precipitation values from the FF ensemble are broadly consistent with the historical record which would be expected given the monthly quantile-quantile mapping approach taken to bias-correct precipitation in the development of the FF scenarios (Newton et al., 2012).

In terms of the frequency and severity of notable precipitation deficits (periods during which the deficit from long term average is greater than 100 mm) this chapter has highlighted that none of the FF precipitation sequences reach a maximum deficit of the magnitude of that associated with the historical 1976 event (325mm), with the two highest deficits present in the FF ensemble 304mm and 292mm. In general there appears to be a tendency for the FF ensemble members to be wetter than the historical record, both in terms of the mean annual precipitation and the mean maximum deficits record during the notably dry periods, with the exception of scenario 'afixl' which contains a dry event that is the closest, in terms of magnitude of precipitation deficit, to that recorded during the 1976 period in the historical record.

For PET, the projections typically show greater variation for each month relative to the historical record than for precipitation, with the FF sequences generally exhibiting lower values in spring and higher values from August through to November compared to the historical record. The research presented here shows significant variation between ensemble members in their estimates of PET and some members can be seen, on average, to be both significantly wetter and

exhibit significantly higher PET rates than the historical record (e.g. member 'afixh') which is likely to be of significance when looking at low flows in a largely permeable catchment.

This chapter also compared river flows between the historical record and the equivalent period of the transient FF ensemble members via inspection of the Flow Duration Curves (FDCs) and changes in Mean Summer Flows (MSFs) for low flow years. The FDCs highlight particularly noticeable divergences for some FF projections, with ensemble member 'afixh' resulting in a FDC that is very different to that generated using the historical climatology.

Analysis of low flow sequences, as measured by looking at MSFs, showed that none of the FF ensemble members contain periods that are as severe as the historical record, with some FFF scenarios not demonstrating any MSF values as severe as the 5<sup>th</sup> most severe event in the historical record (e.g. 'afixh', 'afixc' 'afixo'). This may suggest that the natural climate variability, and the presence of persistent dry periods, within the FF ensemble may be limited which is of particular concern in terms of their application for water resources planning, with member 'afixh' being a particular obvious example.

Given that the underlying climate model runs also underpin the UKCP09 projections, this may raise further concerns as to the adequacy of the underlying models in terms of reproducing the events of interest to water resource planners despite these being the climate products recommended to water resource planners for taking into account climate change (Environment Agency, 2012). This reinforces concerns (Goodess, 2013; Rocheta et al., 2014; Watts et al., 2015) that the outputs from the climate models used to develop the UKCP09 and FF projections have limitations in reproducing the blocking patterns that are significant in the formation of extended droughts.

As outlined in section 2.2.1.1, current climate models typically underestimate the occurrence of blocking, in particular in the Euro-Atlantic sector (Scaife et al., 2011) and this is likely to remain an issue in the near term given that the CMIP5 GCMs, that will underpin the next set of climate projections for the UK (UKCP18), continue to underestimate the observed blocking frequency over Europe (Anstey et al., 2012). This limitation is particularly significant to the water resources planner as it is water resource system's resilience to droughts which drives the planning process and the investment in supply-side and demand-site interventions.

#### *Changes in climate variables under the UKCP09 and Future Flows projections*

The FF and UKCP09 projections suggest a tendency towards hotter, drier summers and wetter winters in the future along with a potential delay in autumn/winter precipitation to replenish soil

moisture stores and recharge aquifers. The monthly changes in precipitation and PET suggested by the FF ensemble members lie within the ranges covered by the UKCP09 projections which are consistent with another study which has compared the two related climate products (Prudhomme et al., 2012).

The UKCP09 ensemble suggests changes in summer (June to August) precipitation, by the 2080s under the Medium Emission scenario, varying from a reduction of 65% to an increase of 37%, with changes in summer temperature ranging from +1 to +9 degrees Celsius. The FF projections suggest mean precipitation reductions ranging from a reduction of 63% to an increase of 47% for specific summer months by the 2080s. These large ranges highlight the very high uncertainty associated with the climate projections currently available to water resources planners in the UK.

#### *Assessing the impacts of climate change on river flows using the UKCP09 and Future Flows projections*

Using the MSME has provided a means of comparing and considering the relative sensitivity of model results to different uncertainties at multiple locations as part of a climate change impact assessment. It is recognised such sensitivities need to be treated with caution as they are conditional on the assumptions made as to the relative skill of each hydrological model and climate projection. Further, the analysis presented here implicitly assumes that the MSME members remain as valid under a future climate as under the observed climate record to which they have been conditioned (i.e. their predictive skill remains unchanged through the future time horizons considered).

A comparison of the range of predicted impacts on flows under an ensemble of climate projections and hydrological models highlights that both the uncertainty from the hydrological modelling and that present within the climate projections are significant, with the latter appearing to be providing a larger contribution to the overall uncertainty (greater variation between climate scenarios for the same hydrological model than between parameter sets for the same climate scenario).

The hydrological modelling using the UKCP09 scenarios presented in this chapter has demonstrated that the Thames catchment may be significantly impacted by such changes in climate, with all catchments studied showing significant decreases in low flows by the 2080s (under a Medium Emission scenario). In terms of a trajectory of impacts over time, the UKCP09 projections suggest increasing reductions on low to medium flows through to the 2080s, relative to a 1961 to 1990 baseline.

The evidence presented in this chapter has also demonstrated the potentially significant influence that the method used to estimate PET may have on the predicted impacts of climate change on river flows for all the time-horizons considered (e.g. see Figure 5-13). Using the Oudin (2005) temperature based method, the central estimate of reduction by the 2080s in Q95 at Kingston is 19.1% compared to 22.3% using the modified Penman-Monteith method (Prudhomme & Williamson, 2013). The difference in the estimated impacts on river flows due to PET method varies across the sampled climate scenarios. This difference reflects the relative contribution from the changes in precipitation and PET with the largest difference due to the PET method under those scenarios with the greatest predicted increase in summer PET. The significance of the method used is further emphasised when examining the results for the perturbed 1976 flow sequence. It should also be noted that it is actual evapotranspiration, which often occurs at less than the potential rate due to crop stress arising from limited soil moisture during the low flow periods (e.g. Rushton et al., 2006; Zhao et al., 2013), that impacts water availability and this will mitigate some of the differences arising from the choice of PET calculation method. These findings suggest that further research is needed to better understand and estimate changes in PET under a future climate along with how hydrological models calculate AET during droughts, with largely permeable catchments, such as the Thames, likely to be particularly sensitive.

*Climate change projections for use by water resource planners, as recommended by current industry guidance*

The research presented in this chapter also highlights a significant challenge in how the two approaches used in the latest WRMPs for assessing climate change allow a water resource planner to consider droughts, the events that typically 'drive' the future management of water resource systems and initiate potentially large investments in infrastructure. The research has demonstrated that both sources of climate change projections have significant limitations for the water resources planner who may want to understand a system's vulnerability to a range of plausible droughts of different intensity, duration and sequencing, and their likelihood of occurrence, representative of a future climate.

The UKCP09 projections, through the perturbation of historical rainfall and PET, result in the system being tested against droughts more intense than those in the baseline record but such an approach makes the fundamental assumption that change in extreme drought events will be similar in magnitude to the changes in mean climate on which the UKCP09 projections have been conditioned. Unlike the FF ensemble, the sequence of droughts is preserved under the UKCP09 climate projections and therefore water resource systems are not tested under the types of droughts not already experienced and to which they may also be vulnerable.

The transient FF scenarios do allow ‘what-if’ scenarios to be undertaken, testing systems under potentially severe drought events not already experienced, but don’t provide any associated probabilities as to their likelihood. Two potentially severe events contained within the FF ensemble were examined and demonstrated that they are potential sequences to which the Thames system may be vulnerable. These events also highlighted that the use of a simple metric such as MSF may not always be a good measure of the potential severity of a particular drought where, for example, low flows persist into the autumn and winter period. However, there is still the underlying concern, reported above, as to whether the magnitude and frequency of droughts in the transient FF projections are appropriate and representative of a future climate and caution is advised in their use in water resources studies such as that published by Huskova et al. (2016).

#### *Assessing the impacts of climate change for water resources in London*

The hydrological model ensemble used here to assess the impacts of climate change, along with the UKCP09 climate change scenarios sampled and examined in this chapter, are taken forward in Chapter 6 to assess the potential impact of different sources of uncertainty on water resources using a water resources system model of London to generate decision-relevant performance metrics.

## Chapter 6 Future reliability of London's public water supplies

### 6.1 Introduction

Water companies in the UK have a statutory responsibility (Water Act, 2003) to produce a plan, revised every 5 years, showing how they intend to maintain the balance between supply and demand for water into the future (considering, as a minimum, a 25 year planning horizon). In England and Wales these plans are known as Water Resources Management Plans (WRMPs). These plans explicitly consider demand-side and supply-side pressures, including population growth and climate change, and describe how the water company intends to secure a sustainable balance between the supply and demand for water to deliver a specified reliability of supplies to customers. These WRMPs are produced at a Water Resource Zone (WRZ) scale. The Environment Agency defines a WRZ as the largest possible zone in which customers share the same risk of a resource shortfall (Environment Agency, 2012).

In developing their WRMPs (e.g. Thames Water, 2014), water resource system models are typically used by water companies. These models use inflow time-series reflecting the availability of supplies, information regarding the water resources infrastructure and profiles of water demand representing intra-year variability. Typically, only a single hydrological model realisation, along with a single method for estimating PET, is used to provide each source's inflows. It is also recognised that there is trade-off between the reliability of supplies and the associated costs of their provision and through consultation with stakeholders, water companies define the planned frequency of customer demand restrictions, for shortages of varying severity, which are known as Levels of Service (LoS).

### 6.2 Objectives

This chapter takes the findings from chapters 4 and 5 forward to consider the influence of hydrological modelling and climate change uncertainty on the future performance of London's public water supply system. This assessment is quantified using the typical terminology of the UK water resources planner and the supply-side impacts are compared to a range of projected demand-side pressures for London. The objective is to provide evidence as to the relative significance of the uncertainty associated with hydrological model parameterisation, the method used to estimate PET and the predicted impacts of climate change on water resource system performance, looking further into the future than the latest WRMPs. The specific research questions posed in this chapter are:

**RQ3.** How do the uncertainties associated with hydrological modelling, estimating PET and climate change all compare in terms of their potential influence on water resources

planning, and how do they compare to demand-side uncertainties? How can the relative significance of different uncertainty sources be quantified in the language of the water resources planner?

**RQ6.** How significant are the different sources of hydrological modelling uncertainty on the performance of London's water resources system? How do the projected impacts of climate change on London's water supplies compare to those anticipated to be caused by population growth and the resulting changes in the demand for water? How do these vary into the future, considering three future time-periods, the 2030s, 2050s and 2080s?

### 6.3 London's water resources system

This section describes the development of a water resources system model of the London water resource zone (Thames Water, 2014) prior to its use to answer the above research questions.

#### 6.3.1 London's water resource system

Water resources in the Thames basin are used to supply homes in London and other areas with potable water, for industrial abstraction and supporting environmental and navigational flows. Water company surface water abstraction in the London area has increased to support a rapidly growing urban population (and increasing per capita consumption) from around 1,000 MI/d ( $\sim 11.6 \text{ m}^3/\text{s}$ ) in the 1920s to around 2,000 MI/d ( $\sim 23.1 \text{ m}^3/\text{s}$ ) today (HR Wallingford Limited, 2014). A large volume of the water abstracted is returned to the river as treated effluent and effectively recycled, but a proportion of this resource is discharged too far downstream to be recovered.

The amount of water available varies considerably between years and is significantly constrained in drought years (Thames Water, 2014). In order to manage these variations and provide a reliable supply to London, water is stored in reservoirs in West London and the Lee Valley (see Figure 6-1). In addition, a number of strategic water supply schemes have been developed to increase supplies to meet peak demands and mitigate the effects of droughts, such as the desalination plant at Beckton.

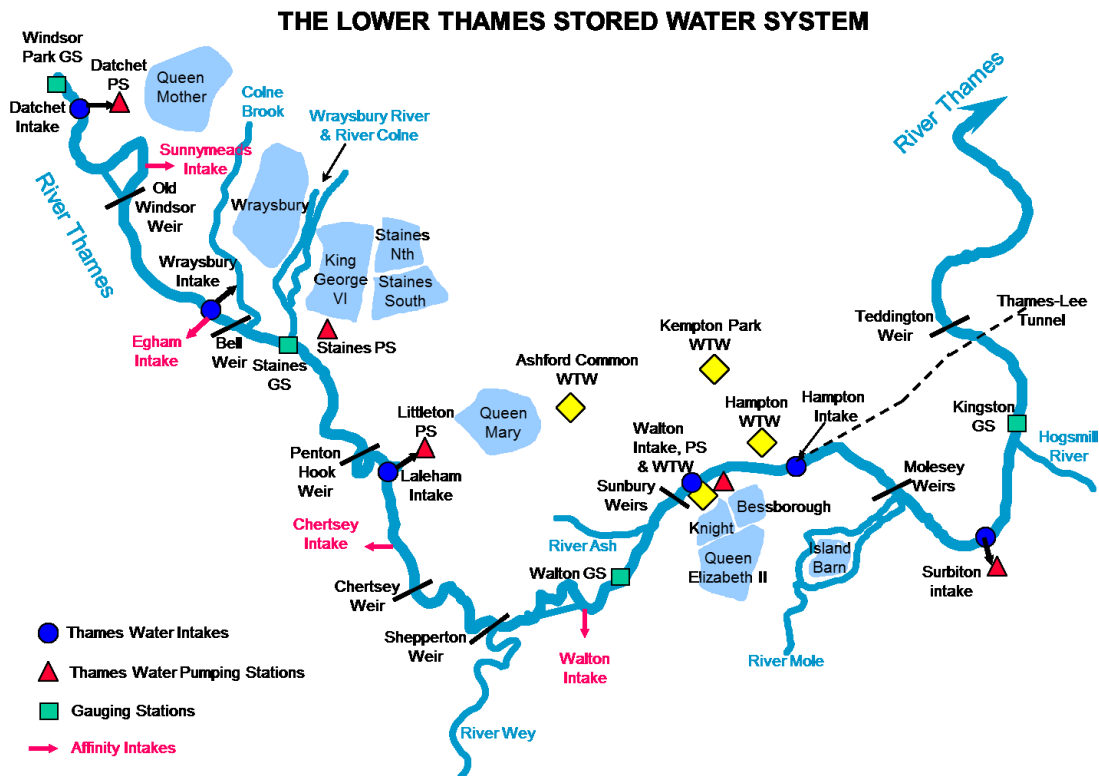


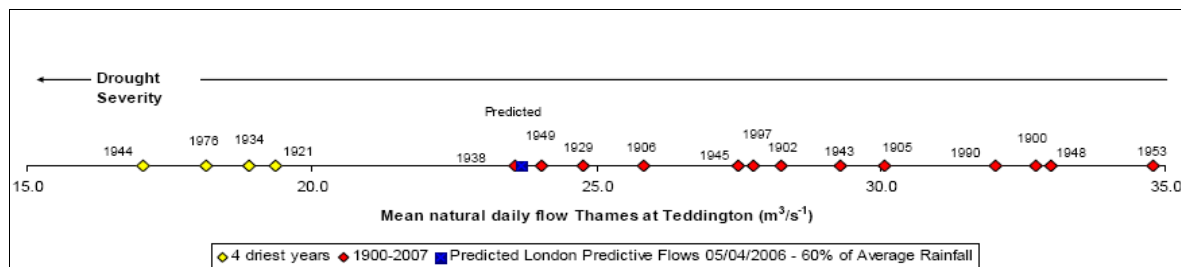
Figure 6-1: Overview of the Lower Thames system including locations of storage and surface water intakes (Thames Water, 2014).

London's water comes from many sources, including boreholes, wells and springs, but most is abstracted from the rivers Thames and Lee and stored in reservoirs before being put into supply. The reservoirs provide a buffer for use in dry periods when abstraction from the rivers is restricted. The quantities that can be abstracted from the rivers depend on the relationship between the quantities stored in the reservoirs, the need to ensure a statutory residual flow over Teddington weir, and the time of year. This is governed by the formal operating agreement between Thames Water and the Environment Agency (EA) under Section 20 of the Water Resources Act 1991, called the Lower Thames Operating Agreement (LTOA) (Thames Water, 2014).

### 6.3.2 Water resources and drought planning in London

Thames Water use Frequency of Occurrence as a way of describing, in statistical terms, the severity of particular droughts in their area. Figure 6-2, taken from Thames Water's Drought Plan, ranks the 20 driest years between 1900 and 2006 based on average naturalised flow (see Section 3.4) in the River Thames (at Teddington) between April to September. This indicator is considered to provide a good approximation to the water resources available in the London water resource zone (Thames Water 2014).





**Figure 6-2: Drought severity as a function of naturalised flow at Teddington (April to September) between 1900 and 2006 (taken from Thames Water’s Final Drought Plan, 2013).**

River flows are also used as drought indicators in Thames Water’s Drought Plan (DP), with specified thresholds on the River Thames, River Wey and River Kennet upon which drought management measures may be activated to mitigate drought conditions. For the London water resource zone, such drought management measures may include activation of the Thames Gateway Water Treatment Works (desalinisation) and the West Berkshire Groundwater Scheme to provide supplementary water during droughts.

As reported in their latest plan, a growing supply-demand deficit is projected for London by 2040 (Thames Water, 2014). Widening supply-demand deficits in London may have a range of consequences such as the reduced reliability of supplies (with more frequent or severe shortages during drought and therefore failure to meet the specified Levels of Service – see Section 6.3.4.4) through to increasing the costs of supplying water. The actual consequences will depend on the size of the deficits and the measures introduced to manage these future risks. A study by NERA (2012) suggested that in London the financial impact of emergency restrictions on water use – which involve rationing the amount of water available for public use – could be in the range of £236 - £329 million every day.

### 6.3.3 Assessing the supply-demand balance

In England and Wales, water companies calculate the supply-demand balance to satisfy the following:

$$\text{Water Available For Use (WAFU)} \geq \text{Distribution Input plus Target Headroom}$$

Water Available For Use (WAFU) is the **Deployable Output** plus bulk supply imports, minus bulk supply exports and minus reductions made for outage allowance and operational losses. In terms of supplies, water companies determine the Deployable Output, typically based on hydrological and water resources modelling, with Deployable Output defined in the Environment Agency’s Water Resources Planning Guidelines (Environment Agency, 2012) as:

*“the output for specified conditions and demands of a commissioned source, group of sources or water resources system as constrained by; hydrological yield, licensed quantities, environment (represented through licence constraints), pumping plant and/or well/aquifer properties, raw water mains and/or aqueducts, transfer and/or output mains, treatment, water quality and levels of service.”*

The Deployable Output (DO) for London is calculated as being equal to the demand that can be met before specified levels of service are breached (see Section 6.3.4.4). Water companies, including Thames Water, calculate climate change impacts on Deployable Output as part of their WRMP.

The total demand for water is presented as the total **Distribution Input**, which is the amount of water that needs to be provided in order to meet customer demands whilst allowing for losses in the system. Distribution Input is typically calculated as the average amount of potable water entering the distribution system within the area of supply (adapted from the definition provided by Rance et al., 2012) with **Target Headroom** representing the minimum buffer that companies should plan to maintain between supply and demand to cater for current and future uncertainties.

#### 6.3.4 Modelling of London’s water resource system

For developing the supply-demand balance for London, Thames Water use a bespoke modelling system called Water Resources Management System (WARMS) which integrates hydrological models (to estimate river flows) and a water resource system model to reflect available system infrastructure and their associated constraints, licences and operational rules and account for within year variations in demand (including the triggering of demand restrictions – see Section 6.3.4.4).

In terms of the water resources model, WARMS represents the water supply area as a network of nodes which include river basins, abstraction points, water treatment works, reservoirs and demand centres connected by links. The main inputs are rainfall, PET, demand profiles, target Levels of Service (see Section 6.3.4.4), demand reduction factors during drought, and licence and infrastructure constraints.

##### 6.3.4.1 The water resources modelling platform, Kestrel - WRM

For the research presented in this chapter, HR Wallingford Limited’s in-house Kestrel – WRM water resources modelling platform, used on a recent study for the Environment Agency exploring the performance of the performance of water supply systems during mild to extreme

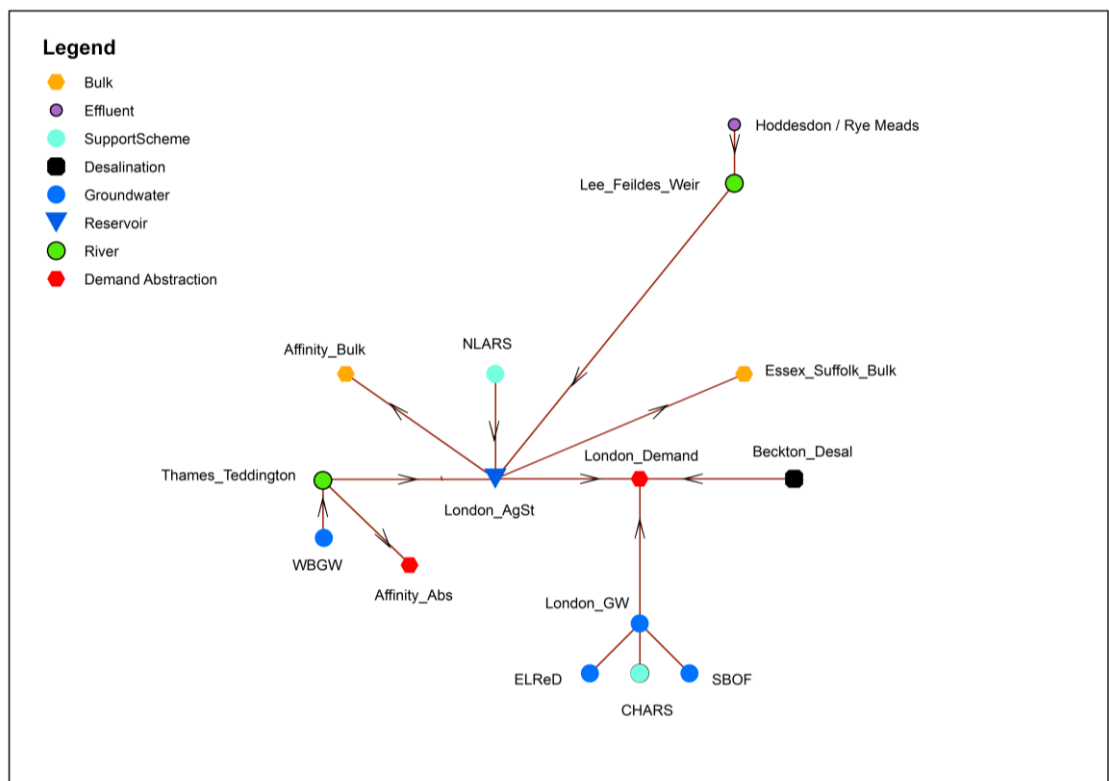
droughts (Environment Agency, 2015), has been used to develop a water resources system model of the London WRZ. Kestrel – WRM is a rule-based, mass balance, node-link computational model similar to models such as IRAS-2010 (Matrosov et al., 2011), *The London Area Rapid Water Resource Model* (LARaWaRM, Walsh et al., 2015) and LANCEMOD (Lopez et al., 2009). Similarly to WARMS, these models conceptualise water resources infrastructure as a network of nodes, such as river intakes, reservoirs and demand centres and links between such nodes. Kestrel – WRM allows operational and demand allocation rules to be implemented which can be based on the state of one or more nodes within the model (e.g. reservoir storage) and other relevant factors such as licence conditions, operational constraints and the time of year.

Within Kestrel – WRM the demand abstraction nodes incorporate flow targets which reflect the demand that needs to be fulfilled at various locations within the system. The Kestrel – WRM model attempts to satisfy these targets at each time-step by requesting water from either supply nodes, storage nodes or transfer links. At each demand node, rules can be specified as to how each resource is accessed and prioritised under different conditions.

Supply nodes incorporate a time-series of available resource (e.g. a river flow series, effluent returns time-series, the specified output from a support scheme) along with the rules related to that resource (such as intake capacities and licence conditions). These rules can be linked to the state of other model nodes (e.g. reservoir storage) within the model. These supply nodes can be linked directly to demand nodes, as described above, or to storage nodes (e.g. reservoirs) or transfer links. These storage nodes in turn have associated operational rules (e.g. storage targets at different times of the year) and constraints which define when, and how, they abstract water from supply nodes (or transfer links, see below). Rules can also be assigned to the storage nodes to trigger releases from storage to another node (e.g. to represent augmentation of river flows). Finally, transfer links allow water to be transported from one node in the system to another node. These can be used to represent bulk transfer agreements (contractual requirements to provide or receive water to/from other water companies) and which can have operational (e.g. capacity) constraints specified dependent upon the state of other nodes in the model.

#### 6.3.4.2 Modelling London's water resource system using Kestrel - WRM

For the purposes of this research, a number of simplifications to the representation of London's water resources infrastructure have been made with a conceptual representation of the water resources system model presented in Figure 6-3. Consequently, whilst the modelling is appropriate for meeting the objectives of this research and quantifying the relative influence of different sources of uncertainty, the absolute values of the supply-demand balance should not necessarily be considered to reflect all the operational constraints (e.g. pump operations, pipe capacities) that are taken into account in the WARMS analysis undertaken by Thames Water in developing their WRMP.



**Figure 6-3: Conceptual representation of the Kestrel – WRM model of the London water resources system.**

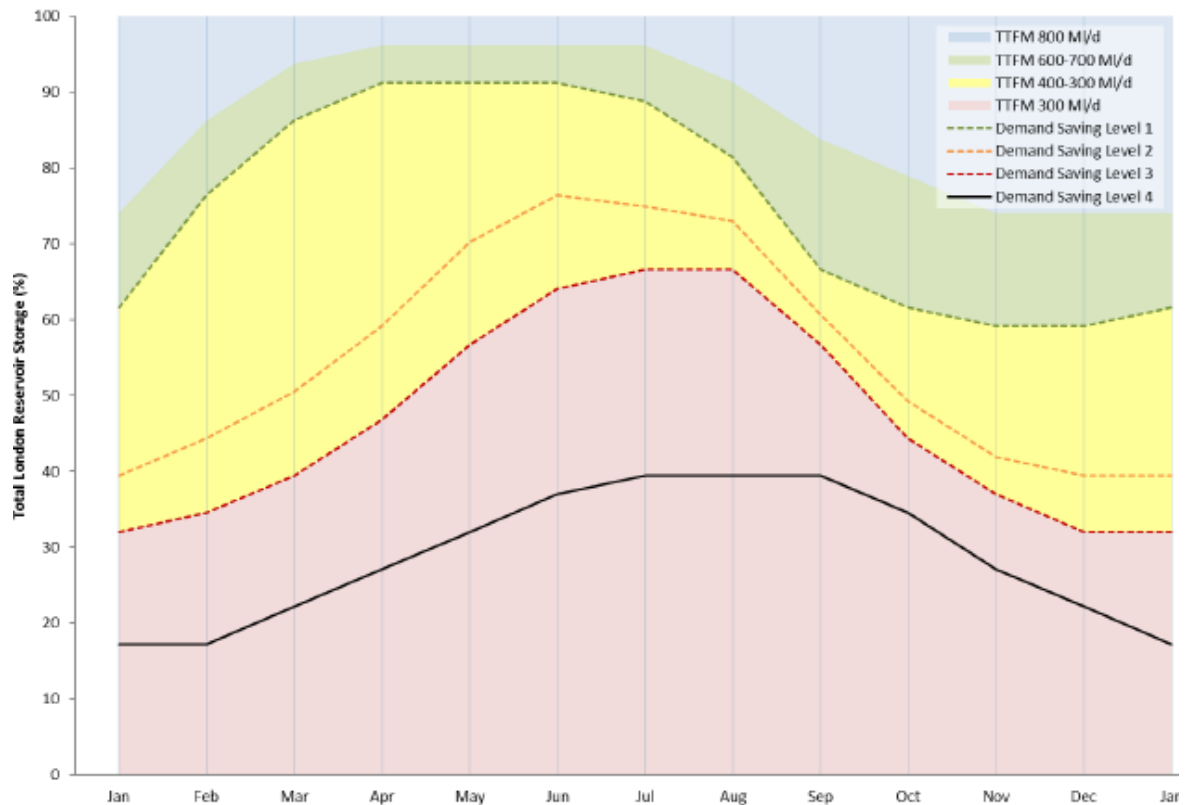
For the modelling presented here, a daily time-step has been used (as is used in WARMS) and is considered appropriate for water resources modelling, particularly where river flows are significantly influenced by baseflow from aquifers and where licence and flow constraints are considered on a daily basis. Borgomeo et al. (2014), in their modelling of London's water supply system using the IRAS-2010 water resources model, adopted a weekly time-step for assessing system performance.

#### 6.3.4.3 Representation of surface water storage in London

The WARMS model aggregates reservoir storage to three lumped volumes – Thames North, Thames South and Lee Valley. For the purposes of the research presented here, the surface water storage available in the London WRZ has been aggregated to a single reservoir node in the Kestrel – WRM model, similar to the approach adopted for Matrosov et al. (2011), Borgomeo et al. (2014) and Walsh et al. (2015).

#### 6.3.4.4 The Lower Thames Operating Agreement (LTOA)

Abstraction from the Thames is managed according to the Lower Thames Operating Agreement (LTOA) with the associated Control Diagram (LTOACD) presented in Figure 6-4. Explicit in the LTOA is the need to maintain a prescribed flow over Teddington Weir which is controlled according to total London storage and the imposition of water use restrictions on customers and the time of year. As London reservoir levels fall, the minimum flow over Teddington Weir (the residual river flow to tidal waters) may be reduced by abstraction in defined bands down to a minimum flow of 300 ML/d ( $3.5 \text{ m}^3/\text{s}$ ). In conjunction with the changing flow constraint, as storage declines the company must apply progressively more intensive demand management measures and restrictions on water use by customers in order to both preserve available storage and mitigate against over abstraction from the River Thames and consequent environmental damage (Thames Water, 2014).



**Figure 6-4: Lower Thames Operating Agreement Control Diagram (LTOACD) showing demand reduction thresholds (Levels 1 to 4 as dashed lines) and minimum environmental flow rates as a function of London's aggregate reservoir storage (LAS).**

Levels of Service (LoS) describe the average frequency that a company will apply restrictions on water use (UKWIR, 2012). As reported in Walsh et al (2015), the failure to meet these Levels of Service can also act as suitable metrics of risk (Hall et al., 2012; Lempert & Groves, 2010) in the robust analysis of water resource systems which have recently become the subject of significant interest within the UK water resources industry (Matrosov et al., 2013; Borgomeo et al., 2014; Anglian Water 2014; Thames Water 2014).

In the case of London, these LoS are related to a range of metrics of which of most importance are the reservoir control curves. As total reservoir levels fall the residual flows over Teddington weir can be reduced and demand restrictions (see Section 6.3.4.8) can be introduced in line with the water company's Levels of Service (LoS). Thames Water's LoS, as implemented through the LTOACD, are as follows (Thames Water, 2014):

- Level 1, intensive media campaign, 1 in 5 years.
- Level 2, enhanced media campaign, sprinkler ban and unattended hosepipe ban, 1 in 10 years.
- Level 3, temporary use ban (TUB), drought direction 2011 (Ordinary Drought Order), 1 in 20 years.
- Level 4, rota cuts and standpipes (Emergency Drought Order), "never" [this use of "never" is as reported in Thames Water's WRMP (2014) to reflect that the contemporary system is designed such that the Level 4 threshold would not be breached if subjected to the historical climate record (1920 to 2010)].

#### 6.3.4.5 River flows and groundwater

The key river sources for considering the supply-demand balance for London are the River Thames and River Lee. In terms of hydrology, Thames Water's WARMS system uses a form of the CatchMod rainfall runoff model for both the River Thames and River Lee and details of its application in WARMS can be found in Appendix I of the Thames Water's WRMP (Thames Water, 2014).

The river flows measured in the Thames basin include the impacts from artificial influences such as abstractions, process water returns and effluent discharges and are also affected by river regulation for navigation. Flows are measured by the Environment Agency and the observed flow record is considered to include all upstream influences. The Environment Agency produce a 'natural flow' record, which is partially naturalised by adding back the abstraction by Thames Water and Affinity Water in the Lower Thames (at Datchet, Staines, Littleton, Walton, Hampton, Surbiton by Thames Water plus abstractions made by Affinity Water at Sunnymeads, Egham, Chertsey and Walton – see Figure 6-1) and is therefore considered to represent the flow available to Thames Water (and Affinity Water) in the Lower Thames. This is the record to which the WARMS hydrology is calibrated and was used in developing the hydrological model ensemble in this research (see Chapter 4).

For the water resources modelling presented here, the abstractions for supplying the London aggregated reservoir storage from the lower Thames are aggregated to a single supply node in the Kestrel – WRM model, with these abstractions constrained by both daily and annual abstraction licences as well as the Lower Thames Operating Agreement (LTOA).

Thames Water abstract at several locations on the River Lee and these abstractions are aggregated to a single supply node in the system model. A licence constraint (commonly known as a 'hands-off' flow) is applied as well as the daily and annual licences in the water resources modelling.

For the groundwater resource, detailed groundwater modelling was outside the scope of this research. Consequently, for the water resources modelling undertaken here, similar to Thames Water's WARMS modelling and the research by Borgomeo et al., (2014), it was assumed that the groundwater resource was equal to the dry year Deployable Output (DO). This has been modelled as a fixed supply node in the Kestrel – WRM model with the maximum rate at which groundwater sources can supply water through a dry period set at that reported by Thames Water (Thames Water, 2014).

#### 6.3.4.6 Abstractions, discharges and treatment losses

At Teddington, an abstraction point (demand node) has been added to the water resources model to reflect Affinity Water's abstractions from the Lower Thames. The abstraction profile has been set to represent the consumptive effect of the abstractions and discharges by Affinity Water, similar to the approach adopted in WARMS (Thames Water, 2014).



Thames Water also provides a bulk supply of raw water to Essex and Suffolk Water and also has to make supplies available to Affinity Water for outage events such as pollution incidents on an emergency basis and as sweetening flow in the connecting tunnel to their treatment works. These influences are taken into account in the water resources modelling using transfer links adopting similar assumptions to those used by Thames Water in their WARMS modelling.

The flows at Feildes Weir on the River Lee are supplemented by effluent returns from Rye Meads treatment works using a similar monthly profile as used by Thames Water in their WARMS modelling (Thames Water, 2014). This is modelled as an effluent supply node in the Kestrel – WRM model.

#### 6.3.4.7 Strategic schemes

There are several strategic water resources schemes in the Lower Thames that are switched on when rivers flows or reservoir levels fall below specified LTOACD threshold levels. These have been modelled as supply nodes in the Kestrel – WRM model. The main strategic scheme is the Gateway desalinisation plant, which can provide up to 150 MI/d ( $1.7 \text{ m}^3/\text{s}$ ) of additional resource during drought conditions. This is triggered when the rolling 10-day average River Thames flow at Teddington remains at or below 3,000 MI/d ( $34.7 \text{ m}^3/\text{s}$ ). Supporting flows of 25 MI/d ( $0.3 \text{ m}^3/\text{s}$ ) are provided between April and June.

Other strategic schemes include the West Berkshire Groundwater Scheme (WBGs), the North London Artificial Recharge Scheme (NLARS), Chingford Artificial Recharge Scheme (CHARS) and the Hoddesdon Transfer Scheme. The WBGs release water into the Thames when Level 2 on the LTOACD is crossed, with the release and replenishment profiles used by Thames Water adopted here. NLARS and CHARS are both activated when the when Level 1 on the LTOACD is crossed and is assumed by Thames Water to be able to provide additional resource up to a fixed maximum value for up to 16 months (Thames Water, 2014).

The Hoddesdon Transfer scheme provides additional effluent returns on the River Lee and the profile assumed in WARMS has been used in the modelling presented here.

#### 6.3.4.8 Demand profiles and restrictions

The WARMS modelling uses a standard demand profile with peak demand in August and minimum demand in April each year (Figure 6-5) and this has been adopted for the water resources system modelling presented in this thesis. The estimated impacts of demand restrictions assumed during droughts have been taken from Thames Water's Drought Plan (Thames Water, 2013) and are summarised in Table 6-1.

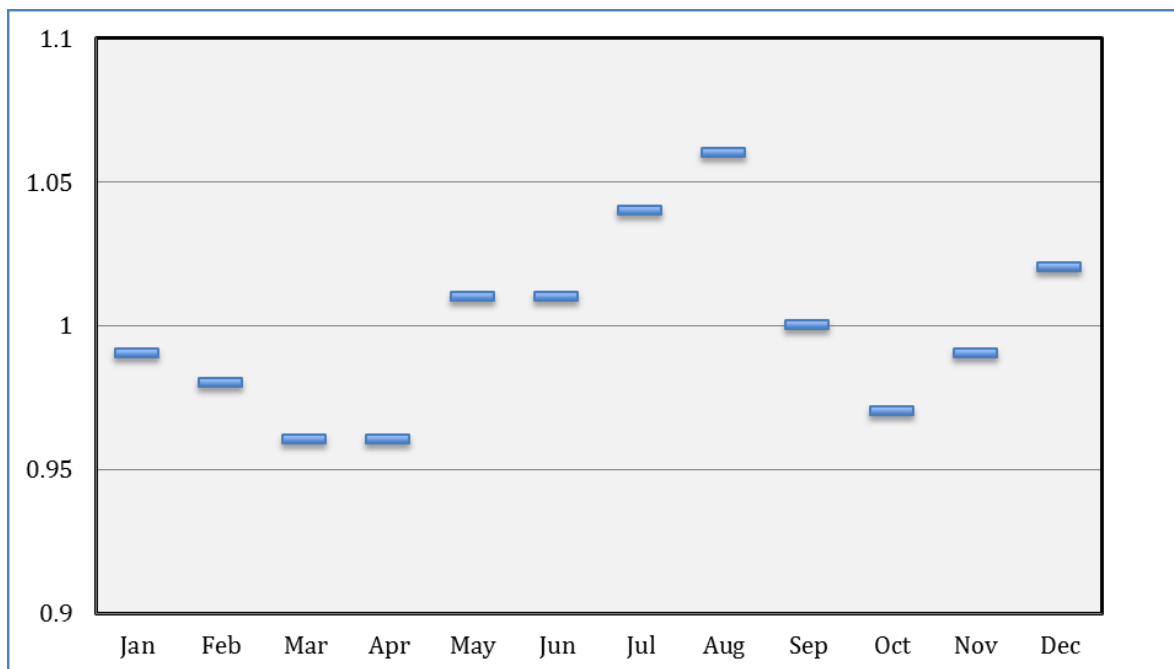


Figure 6-5: Monthly demand factors for supply-demand modelling (Thames Water, 2014).

Table 6-1: Estimates for the impacts of demand reduction for London WRZ, expressed as percentages (Thames Water Drought Plan, 2013).

Threshold	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Level 1	1.1	1.2	1.2	1.3	1.5	2.2	2.2	2.1	1.3	1.2	1.2	1.2
Level 2	2.2	2.2	2.3	3.2	3.8	7.5	7.9	7.1	3.2	2.6	2.4	2.3
Cumulative Levels 1-2	3.4	3.4	3.4	4.5	5.3	9.7	10.1	9.2	4.5	3.8	3.6	3.5
Level 3a	0.7	0.7	0.7	1.1	1.4	2.8	3.0	2.7	1.1	0.9	0.8	0.8
Level 3b	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4
Cumulative Levels 1-3	5.4	5.5	5.5	7.0	8.0	13.9	14.5	13.2	7.0	6.1	5.8	5.6
Level 4	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0
Total Savings Levels 1-4	23.4	23.5	23.5	25.0	26.0	31.9	32.5	31.2	25.0	24.1	23.8	23.6

## 6.4 Baseline modelling of the London water resource zone

The first part of this section presents a comparison of the performance of the Kestrel – WRM model with the outputs from the WARMS modelling presented by Thames Water in their latest WRMP. Following this, prior to considering the supply-demand balance into the future, the reservoir drawdown, and the associated LTOACD level, for the ‘baseline’ period using the model ensemble described in Chapter 4 is compared to the same profile generated using the historical flow record. The period used for this comparison is from 1961 to 2008, as used for the hydrological modelling presented in Chapter 4.

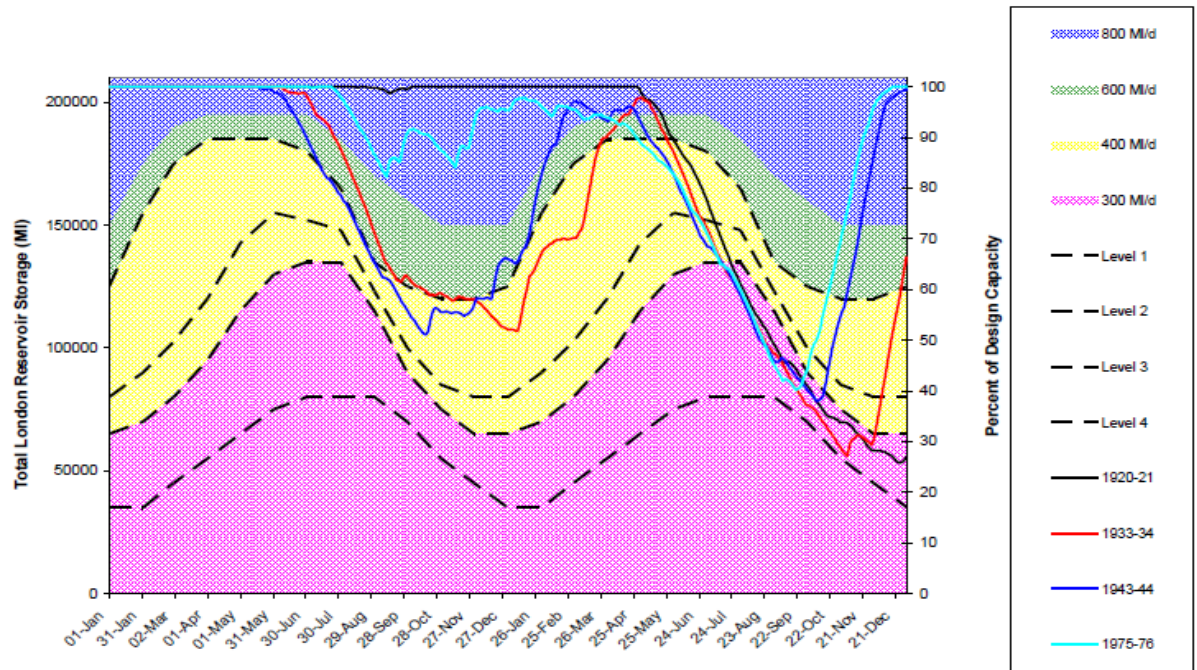
For this initial baseline modelling, the flows in the River Lee were based on the flow record used by Thames Water in their WARMS modelling. Following this, a hydrological model ensemble of the Lee (described in Section 6.4.3) was also developed, with its performance assessed against the historical record, to enable the climate change impacts of flows in the River Lee to be considered coherently with those in the River Thames (see Section 6.4.3).

### 6.4.1 *Performance of Kestrel – WRM compared to the WARMS system*

The objective of this chapter is to examine the potential significance of different components of hydrological modelling uncertainty on the planning process, rather than to repeat Thames Water’s WRMP, or parts thereof such as the supply-demand balance.

London, and the Thames catchment as a whole, has undergone a significant number of changes since the 1960s, including the demand for water and the provision of public water supplies. This makes it difficult to validate the performance of a model against the historical record in terms of, for example, reproducing reservoir drawdowns. However, it is useful to compare the outputs from the water resources model used here with published outputs from Thames Water’s WARMS system under similar conditions (Thames Water, 2009; 2014).

Figure 6-6 is taken from TWUL’s 2014 WRMP and includes a modelled drawdown of the London aggregate reservoir storage for the contemporary water resource system during the 1975 to 1976 climate sequence, taking into account the estimated impact of demand restrictions (see Section 6.3.4.8) and the operation of strategic schemes that would be expected to be employed if such an event was to occur today (see Section 6.3.4.7).

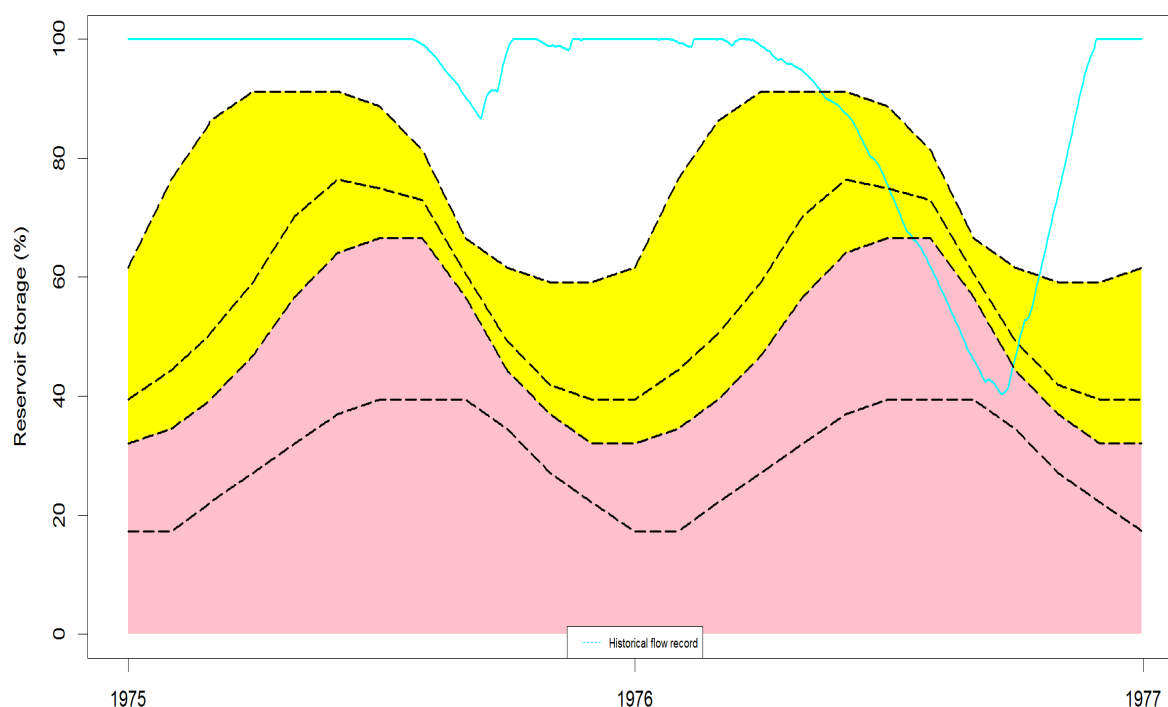


**Figure 6-6: Profile of the LAS during selected periods in the historical climate record using the WRMP demand profile and existing water resources system modelled using the WARMS model (Thames Water, 2014). The black dashed lines and coloured shading reflect the LTOACD (see Figure 6-4).**

Figure 6-7 provides the equivalent plot using the Kestrel – WRM model for London, using the same colour scheme for the four LTOACD levels. The plot demonstrates that during the summer of 1976, the Kestrel – WRM model is able to reproduce the reservoir drawdown reasonably well, albeit with some differences in the lead-up to this event.

The differences between the reservoir drawdown in WARMS and Kestrel – WRM is due to differences in how the historical record is processed and the level of detail adopted in the water resource system modelling. The WARMS system include a more detailed representation of the London Water Resource system compared to the Kestrel – WRM model and includes constraints such as pump capacities, explicit representation of the individual storage reservoirs and their connectivity. Further, for the WARMS system, the inflows for the Thames are represented in greater detail rather than being aggregated to a single node as for the Kestrel – WRM model.

To provide the inflows in WARMS, CatchMod is calibrated using gauged river catchments which are then also translated to a large number of smaller ungauged catchments. This means that flows are generated through a combination of rainfall-runoff modelling and translation of flows between sites. However, given the close correspondence between the drawdown sequences, the Kestrel – WRM model is considered to provide a reasonable basis for evaluating the relative significance of different sources of uncertainty on the water resources planning process.



**Figure 6-7: Profile of the LAS during the 1975 to 1976 period using the Kestrel - WRM model (solid cyan line). The black dashed lines and red/yellow shading reflect the LTOACD (see Figure 6-4). The yellow shaded areas are Levels 1 and 2, with the red areas Levels 3 and 4 – as used in Figure 6-6.**

#### 6.4.2 Baseline water resources modelling using the MSME

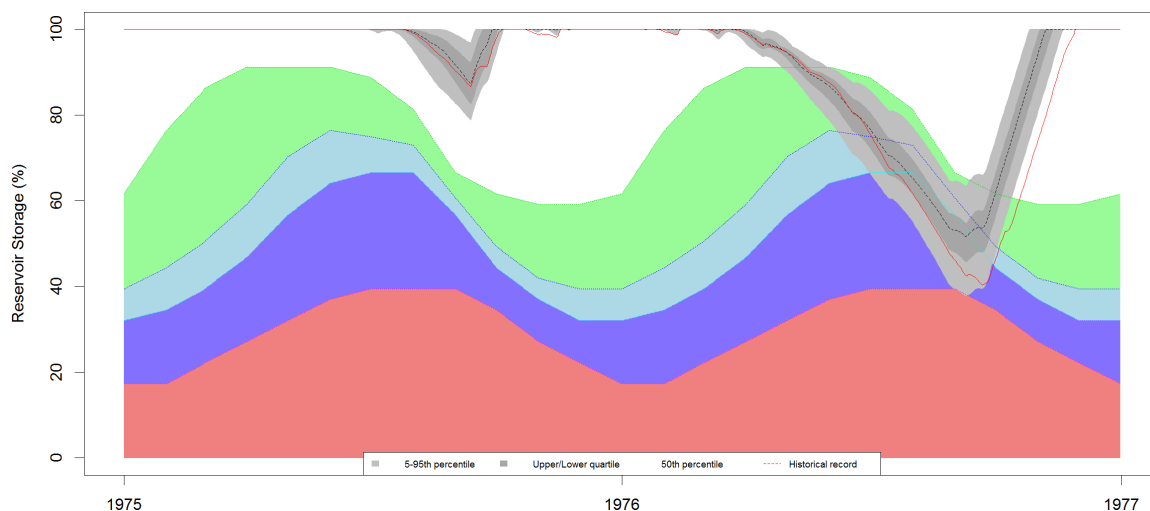
This section compares the performance of the water resources system during the baseline period (1961 and 2008) using the Multi-Site Model Ensemble (MSME) described in Section 4.3. This comparison considers reservoir drawdown and the associated LoS against which water companies plan.

##### 6.4.2.1 Performance of the water resources model driven by the hydrological model ensemble

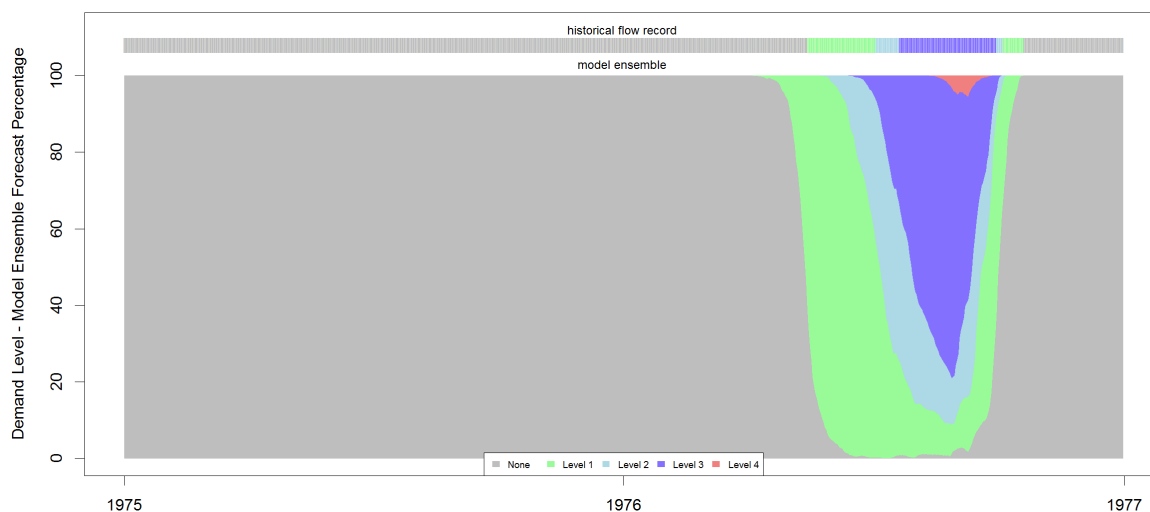
Figure 6-8 and Figure 6-9 present the distribution of the MSME simulations for the London aggregate storage and the associated predicted LTOACD level respectively, compared to the values produced using the historical flow record, for the 1975 to 1976 drought period. The hindcast is produced by weighting the predicted storage each day from each MSME ensemble member using their associated conditional probabilities (see Section 4.3).

The MSME simulations exhibits two key differences to the historical flow record. The MSME prediction of the magnitude of the reservoir drawdown is not as large as for the historical flow record and this is accompanied with the predicted timing of the maximum deficit being slightly earlier than using the historical flow record. Table 6-2 summarises the performance of the system in terms of the number of days at each LTOACD level.

These results for reservoir drawdown in the water resources model are consistent with the river flow simulations presented in Section 4.4.1.2 using the MSME which were generally higher during 1976 than in the historical record. The MSME simulations also predicted a recovery in river flows earlier than present in the historical record and these tendencies lead to an under-estimation of reservoir drawdown during this event.



**Figure 6-8: Profile of LAS during the 1975 to 1976 period comparing the MSME prediction (black line and grey shading) and the profile generated using the historical flow record (red line). The coloured areas reflect the LTOACD (Levels 1 to 4, green to red shading – see Figure 6-4).**

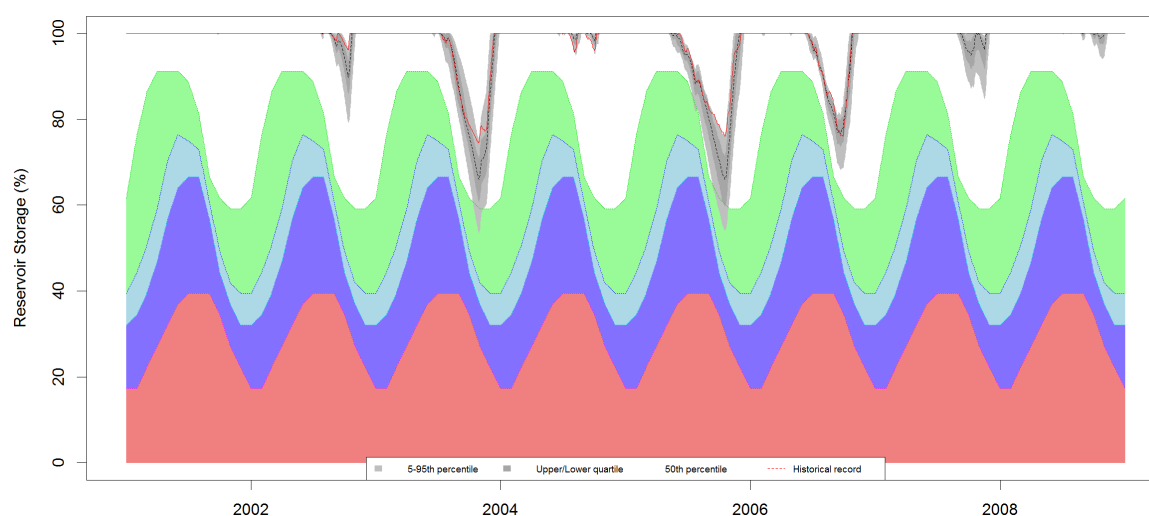


**Figure 6-9: Forecast LTOA levels during the 1975 to 1976 period comparing the MSME prediction (bottom plot) and the profile generated using the historical flow record (top plot). The coloured areas reflect the LTOACD (Levels 1 to 4, green to red shading – see Figure 6-4).**

**Table 6-2: Number of days at each LTOACD Level predicted using the MSME compared to those generated using the historical flow sequence, during 1975 to 1976, at Teddington.**

LTOACD Level	Historical flow record	MSME prediction (percentiles)				
		P5	P25	P50	P75	P95
0	573	565	579	589	597	621
1	65	50	57	65	76	112
2	22	0	24	29	36	49
3	71	0	10	45	68	82
4	0	0	0	0	0	6

During the 1961 to 2008 period, only the summer of 1976 had a significant impact on LoS (see Figure 6-10) and highlights the limited number of historical drought events with which to assess the resilience of a water resources system. Figure 6-10 examines the historical and MSME predicted reservoir drawdown between 2001 and 2008, a period identified in Chapter 4 as contrasting with the 1971 to 1980 period in that the MSME predicted lower flows than the historical record. The impact of this on reservoir drawdown is readily apparent, although in these cases the impacts on the LoS were modest as the severity of the low flow periods were much less than the 1976 event.



**Figure 6-10: Profile of the LAS during the 2001 to 2008 period comparing the MSME predictions (black line and grey shading) and the profile generated using the historical flow record (red line). The coloured areas reflect the LTOACD (Levels 1 to 4, green to red shading – see Figure 6-4).**

#### 6.4.2.2 Assessing the potential influence of method of estimating PET

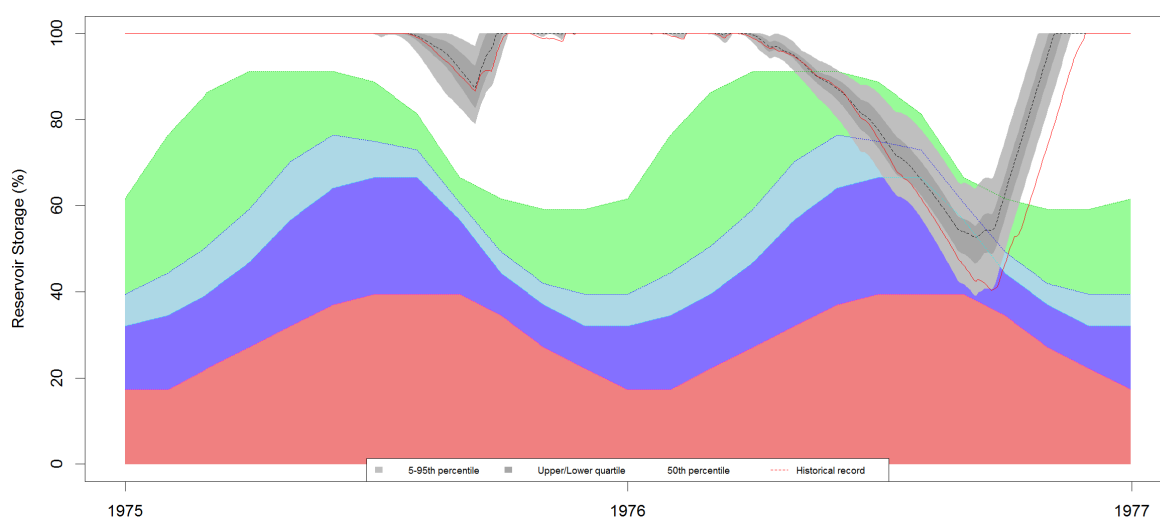
The influence of the method used for estimating historical PET on predicted reservoir levels, and levels of service, was also considered. It can be seen in Figure 6-11 and Figure 6-12 that the predicted reservoir drawdown during 1975 and 1976 using the MSME members developed using the Oudin PET estimation method (known as MSME-OU) and the MSME members developed

using the modified Penman-Monteith estimation method (MSME-PM) separately are similar, with the MSME-PM simulations slightly drier. Table 6-3 summarises the number of days at each LTOACD level using the model ensembles for each PET method along with the same values using the two PET ensembles combined (MSME). This confirms the slightly drier results using the modified Penman-Monteith method, in terms of the number of days at LTOACD levels 3 and 4.

These outputs emphasise that the choice of method used to determine PET, in conjunction with the model conditioning approach adopted, has only a limited effect on model predictions in terms of reservoir storage during the 1975 to 1976 period. In terms of LTOACD levels, the MSME-PM predicts a higher number of days at LTOACD Level 3 although both ensembles (MSME-PM and MSME-OU) under-estimate the duration of the LTOACD Level 3 transgressions compared to using the historical record. The influence of the PET method used under climate change scenarios is examined later in this chapter.

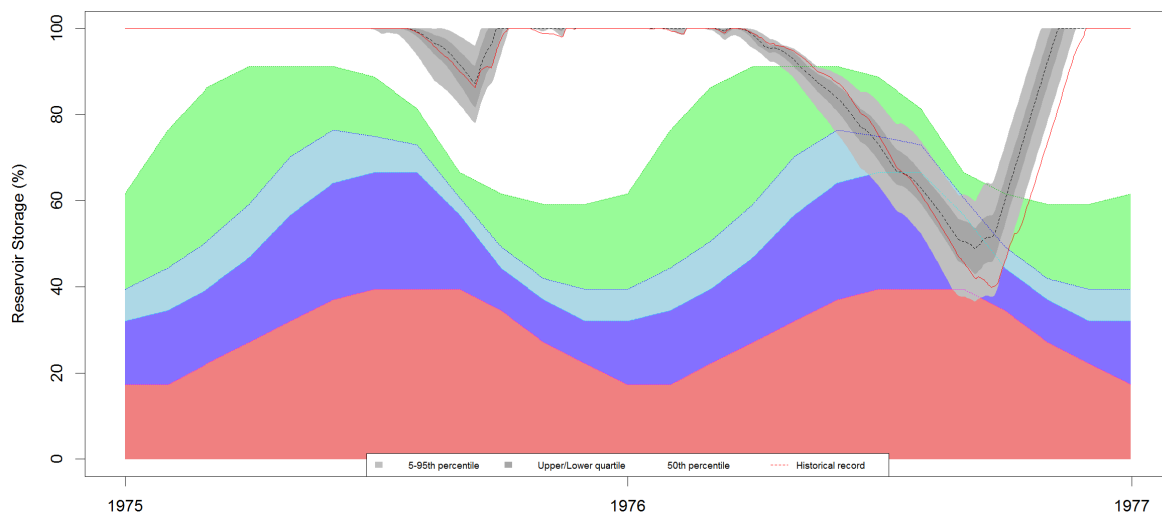
**Table 6-3: Number of days at each LTOACD Level predicted using the MSME compared to those generated using the historical flow sequence, during 1975 to 1976, at Teddington.**

LTOACD Level	Historical flow record	MSME			MSME-OU			MSME-PM		
		P5	P50	P95	P5	P50	P95	P5	P50	P95
0	573	565	589	621	572	591	627	557	581	607
1	65	50	65	112	50	67	113	51	62	106
2	22	0	29	49	0	30	49	14	29	48
3	71	0	45	82	0	41	79	0	59	86
4	0	0	0	6	0	0	0	0	0	19



**Figure 6-11: Profile of the LAS during the 1975 to 1976 period comparing the predictions using the MSME –OU ensemble (black line and grey shading) against the profile generated using the historical flow record (red line). The coloured areas reflect the LTOACD (Levels 1 to 4, green to red shading – see Figure 6-4).**



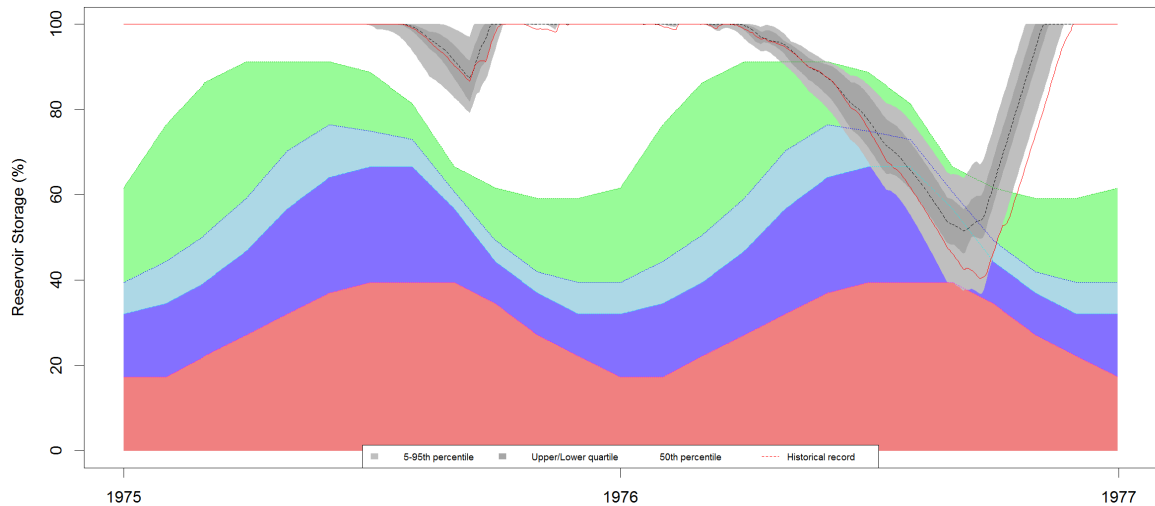


**Figure 6-12: Profile of the LAS during the 1975 to 1976 period comparing the predictions using the MSME-PM ensemble (black line and grey shading) against the profile generated using the historical flow record (red line). The coloured areas reflect the LTOACD (Levels 1 to 4, green to red shading – see Figure 6-4).**

#### 6.4.2.3 Sampling from the hydrological model ensemble

Prior to assessing the projected impacts of climate change on the performance of the water resources system, an approach to reducing the computational demand of combining multiple hydrological model realisations and climate change scenarios (and levels of demand – see Section 6.5) was explored. This was based around running a subset of the MSME that would be representative of the full ensemble but not require all MSME members to be used.

Using the conditional probabilities produced in Chapter 4, samples of 100, 200, 500 and 1,000 were taken and their corresponding outputs from the water resource system modelling compared to the outputs produced using the MSME. The metrics used were visual inspection of reservoir drawdown and the frequency and duration of LTOACD level transgressions. The profile of reservoir storage for a 200-member sub-sample is presented in Figure 6-13, with Table 6-4 summarising the number of days at each LTOACD level for the full ensemble and this 200-member sub-sample respectively. These results demonstrate that the 200-member sample (known as the MSME-200) predicted a very similar frequency and number of days at each LTOACD level (see Section 6.3.4.4) compared to using the MSME and visually a similar reservoir drawdown profile (see Figure 6-8 for the reservoir drawdown using the MSME). Further benefits (as measured by these metrics) were not evident using 500 and 1,000 member ensembles despite the additional computational effort.



**Figure 6-13: Profile of the LAS during the 1975 to 1976 period from the historical climate record. Range of values for reservoir drawdown using the MSME-200 is shown as a black line and grey shading. The storage profile using the historical flow record is shown as a red dashed line for reference. The coloured areas reflect the LTOACD (Levels 1 to 4, green to red shading – see Figure 6-4).**

**Table 6-4: Number of days at each LTOACD Level predicted using the MSME and the 200 member MSME-200, during 1975 to 1976, at Teddington.**

LTOACD Level	Historical flow record	MSME			MSME-200		
		P5	P50	P95	P5	P50	P95
0	573	565	589	621	566	591	626
1	65	50	65	112	51	65	114
2	22	0	29	49	0	29	49
3	71	0	45	82	0	45	83
4	0	0	0	6	0	0	8

#### 6.4.3 Hydrological uncertainty related to flows on the River Lee

Whilst the proportion of the resource abstracted from the River Lee is much lower than that from the River Thames, the Lee still represents an important resource for London with a catchment area around 10% of that of the Thames to Kingston. For the water resource modelling presented previously in this chapter, the flows for the River Lee have been taken from the historical flow record for the River Lee used by Thames Water in their WARMS system.

To enable coherent (with the River Thames) climate change sequences to be available for the River Lee, a hydrological model ensemble was developed using the same approach as for the River Thames, with for the Lee model performance assessed at a single location, Fieldes Weir (NRFA gauge: 38001). Model performance was evaluated using the same six performance metrics used for the Thames for each of the same five time windows. As for the River Thames, both methods of estimating PET were considered and 10,000 parameter sets were sampled, using a

Latin-Hypercube Sampling (McKay et al., 1972) approach. Further details of the methodology can be found in Section 4.3.4.

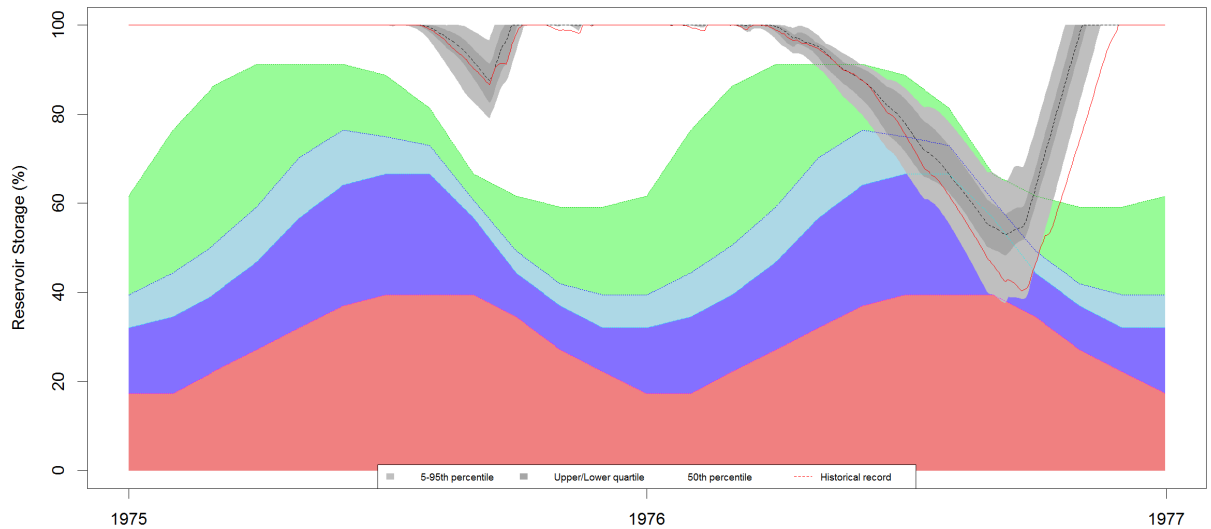
The impact of using this hydrological model ensemble for the River Lee rather than the historical flow record was then explored. For each member of the sub-sample, the flow series for the River Lee was sampled based on their associated conditional probabilities. To generate a conditional probability for each water resources model simulation to cover both rivers, it was necessary to combine the conditional probabilities for the models on the Thames and Lee, with three alternative approaches considered:

- using only the conditional probabilities for the Thames (effectively the results presented in Figure 6-13);
- multiplying the conditional probabilities for the Thames with the conditional probabilities for the Lee;
- adding the conditional probabilities for the Lee to the conditional probabilities for the Thames but with the conditional probabilities weighted by the upstream catchment area.

Based on this, the final approach was considered to best reflect the relative influence of each surface water source with the resulting ensemble known as MSME-200-LEE. A review of the performance of the combined model ensembles during 1976 for this selected approach is presented in Figure 6-14. Table 6-5 summarises the predicted number of days at each LTOACD level compared to using the historical flow records. These outputs demonstrate that MSME-200-LEE simulations, with each member including a sampled flow series from both the Thames and the Lee, provides an acceptable representation of the London Water Resource system for evaluating the performance of London's water resources system under a range of projected climate change scenarios and comparing the relative influence of different sources of uncertainty.

**Table 6-5: Number of days at each LTOACD Level predicted using the MSME-200-LEE during 1975 to 1976.**

LTOACD Level	Historical flow record	MSME-200-LEE		
		P5	P50	P95
0	573	566	591	629
1	65	50	67	110
2	22	0	30	55
3	71	0	37	84
4	0	0	0	5



**Figure 6-14: Profile of the LAS during the 1975 to 1976 period from the historical climate record. The range of values for reservoir drawdown using the MSME-200-LEE is shown as a black line and grey shading. The storage profile using the historical flow record is shown as a red dashed line for reference. The coloured areas reflect the LTOACD (Levels 1 to 4, green to red shading – see Figure 6-4).**

## 6.5 Impacts of climate change on the water resources of London

For assessing the impacts of climate change, the UKCP09 (Murphy et al., 2009b) projections (using the Medium Emission scenario) are used. Similarly to other studies (e.g. Wilby & Harris, 2006; UKWIR, 2007; Prudhomme et al., 2010; Charlton & Arnell, 2014) that have used climate change factors in the context of water resources, this research has used the UKCP09 change factors to perturb the historical climate record prior to being input into the hydrological models.

The climate change impacts considered in this research are based on the 20-member UK climate change scenario sub-samples identified in Section 5.3.1, using two alternative methods for estimating PET. For the water resources modelling presented in this chapter, these 20 climate change scenarios, for each future time-slice, were simulated using the Kestrel – WRM model of London's water resource system, using the MSME-200-LEE.

The period of analysis for assessing the impacts of climate change is the 1961 to 1990 period which is the baseline period used for the UKCP09 climate change projections. The impacts are quantified in terms of the change in the frequency and the length of Level 3 and Level 4 LTOACD transgressions and, later in the chapter, in terms of the reduction in Deployable Output (DO) necessary to mitigate these impacts.

For the baseline 1961 to 1990 period (without climate change impacts) the MSME-200-LEE simulated number of years which experience LTOACD Level 3 and Level 4 breaches, and the total number of days forecast at each of these levels, is presented in Table 6-6 and Table 6-7 respectively. In addition, the simulated estimates using MSME-200-LEE members developed using

the Oudin PET estimation method (known as MSME-200-LEE-OU) and using the MSME-200-LEE members developed using the modified Penman-Monteith PET estimation method (known as MSME-200-LEE-PM) are also separately shown. The impacts of climate change are subsequently considered using these metrics.

**Table 6-6: Simulated number of years at LTOACD Levels 3 and 4 for the baseline (1961 to 1990) period.**

<b>LTOACD Level</b>	<b>MSME-200-LEE</b>	<b>MSME-200-LEE-OU</b>	<b>MSME-200-LEE-PM</b>
3	0.66	0.61	0.78
4	0.06	0.05	0.10

\*calculated as the sum of the number of years at which the specified LTOACD level was the maximum level predicted that year, with each ensemble member total weighted by their conditional probability

**Table 6-7: Simulated number of days at LTOACD Levels 3 and 4 for the baseline (1961 to 1990) period.**

<b>LTOACD Level</b>	<b>MSME-200-LEE</b>			<b>MSME-200-LEE-OU</b>			<b>MSME-200-LEE-PM</b>		
	P5	P50	P95	P5	P50	P95	P5	P50	P95
3	0	37	84	0	11	81	0	68	84
4	0	0	5	0	0	0	0	0	34

\*calculated as the total number of days at each LTOACD level during the 1961 to 1990 period, with each ensemble member total weighted by their conditional probability

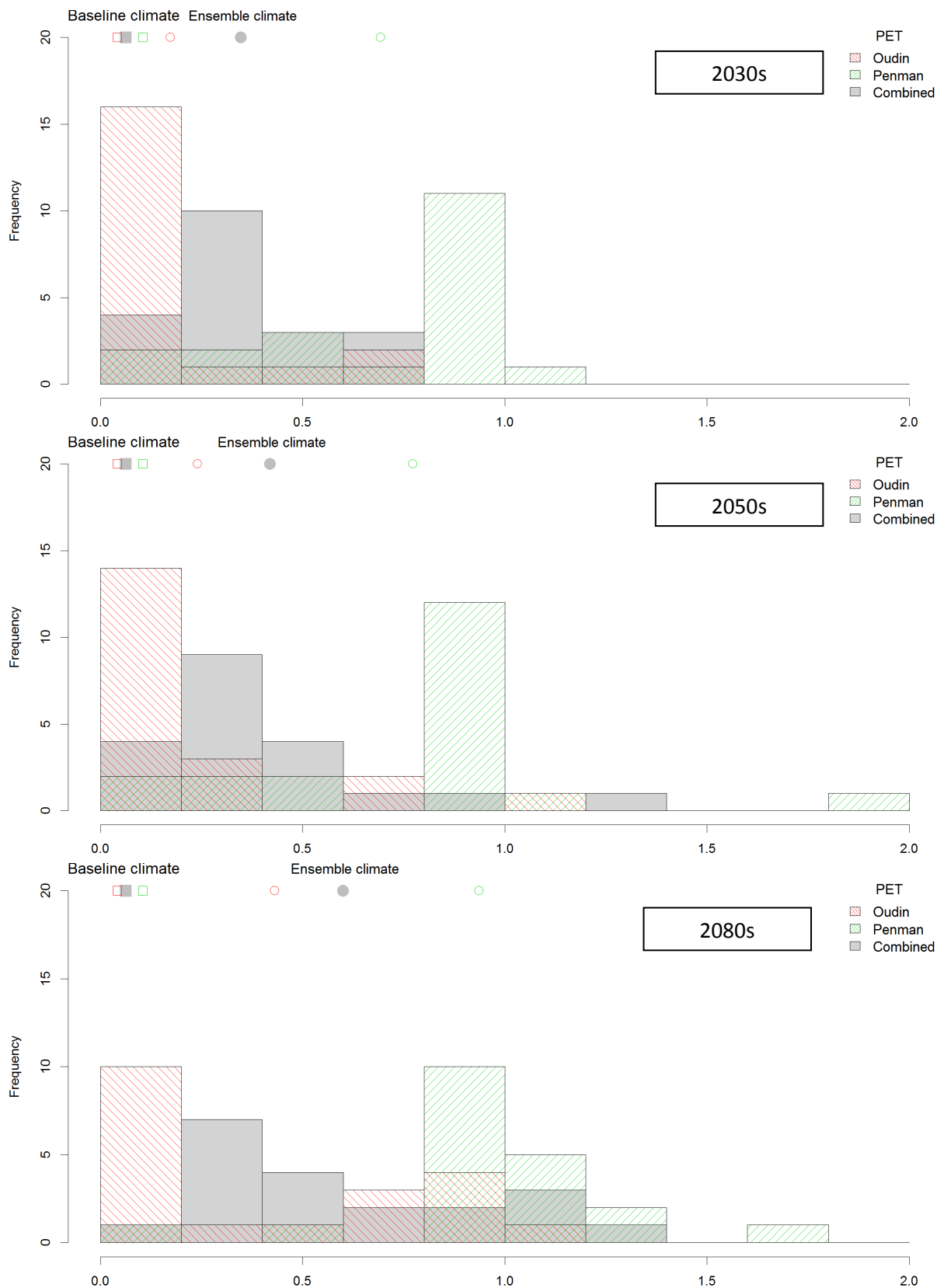
#### 6.5.1 Projected impacts of climate change on Levels of Service

Figure 6-15 presents the distribution of LTOACD Level 4 transgression frequency (the number of years where Level 4 is breached) during the perturbed 1961 to 1990 period for each of the sampled UKCP09 climate scenario from each of the future time-horizons. These distributions are presented separately for both methods of estimating PET (MSME-200-LEE-OU and MSME-200-LEE-PM) and also as a single combined ensemble (MSME-200-LEE) based on their individual conditional probabilities. This water resource system metric is used as a measure of the relative severity of each climate change scenario in terms of the change in *frequency* that impacts that would be faced by customers – in this case rota cuts and standpipes (see Section 6.3.4.4).

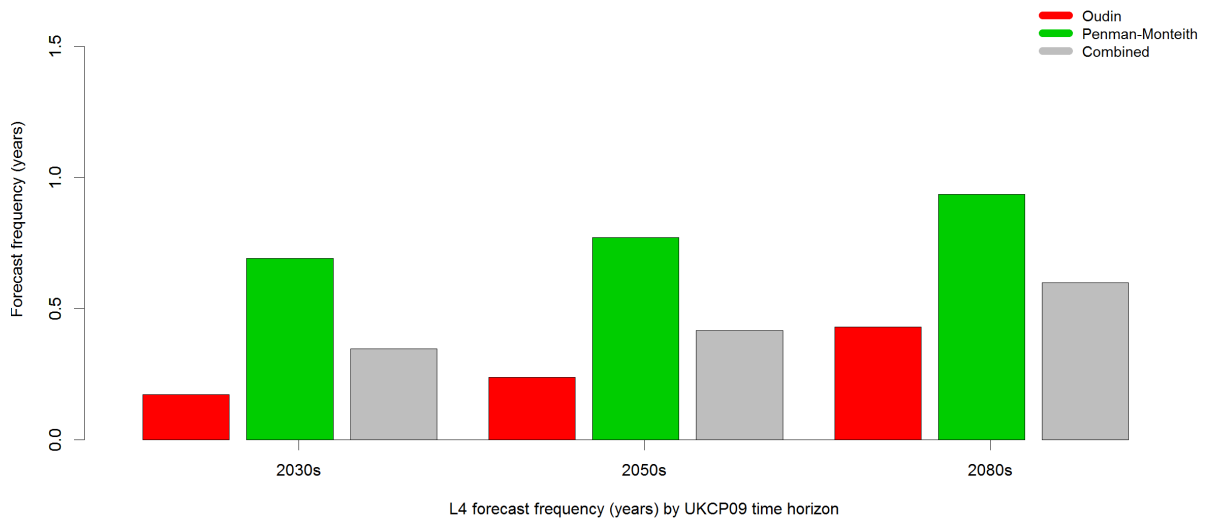
For each of the 20 sampled climate change scenarios for each time horizon, the number of years that each hydrological model ensemble member breaches a LTOACD Level 4 line is calculated. An overall value (for all members of a hydrological model ensemble) for each climate change scenario is then determined by weighting the value of each ensemble member's conditional probabilities as described in Section 6.4.3. The values for each climate scenario are then plotted in the histogram presented in Figure 6-15, with bin widths of 0.2. The small circles above the histograms highlight the overall frequency across the climate change ensemble (considering each climate change scenario as equally valid) and the small squares above the histograms represent

the equivalent frequency derived using the unperturbed climatology (i.e. baseline climatology). Consequently, the impact of the sampled climate change scenarios for each time-horizon can be seen in terms of their relative impacts on LTOACD Level 4 transgressions, a metric of direct interest to water resources planners.

These plots demonstrate a wide range of values across each climate ensemble and, in general, an increasing frequency of LTOACD Level 4 breaches into the future due to climate change. Such a projected trend would be expected based on the river flow results presented in Section 5.4.2 that showed a corresponding trend in reducing flows during low flow periods due to climate change. Despite the limited influence of the method used to estimate PET on the performance during the baseline conditions (see Section 6.4.2.2), these plots also confirm the findings from Figure 5-14 and Figure 5-16 that under a changed climate the method used to estimate PET has a significant impact. The differences in river flow impacts due to climate change manifest themselves into impacts on water resources system performance with the MSME-200-LEE-PM simulations demonstrating much larger impacts in terms of the frequency of LTOACD Level 4 transgressions. This is further evident in Figure 6-16 which summarises the overall ensemble values for each of the three time-horizons considered.



**Figure 6-15: Histogram of the frequency of Level 4 LTOACD transgressions for each UKCP09 climate change scenario sub-sample members for the 2030s (top), 2050s (middle) and 2080s (bottom). The transgressions for the baseline climatology and combined 20-member ensemble are represented by the small circles and small squares above the histograms respectively.**



**Figure 6-16: Forecast frequency of years during which a LTOACD Level 4 breach is realised for different time-horizons and distinguishing between the different approaches to estimating PET.**

Similar tendencies can be observed when looking at a related metric which is the total number of days at LTOACD Level 4 (duration of customer restrictions) during the perturbed 1961 to 1990 periods as shown in Figure 6-17, Figure 6-18 and Figure 6-19 for the 2030s, 2050s and 2080s respectively (note that the small circles are individual values that are more than 1.5 times the inter-quartile range from the extent of the inter-quartile range). This alternative measure also provides an indication as to the relative severity of each climate change scenario in terms of the change in *duration* of the impacts faced by customers in London (see Section 6.3.4.4). These plots also highlight the relative uncertainties associated with the climate change projections relative to the uncertainty associated with the hydrological model parameters suggested by their conditional probabilities. The results indicate that, certainly by the 2080s, the former is larger (as shown by the variation between climate scenarios) but that hydrological parameter uncertainty is still potentially significant (i.e. reflected in the variation within each individual climate scenario). Borgomeo et al. (2014) also reported that hydrological model parameter uncertainty had a significant impact on London’s system performance.



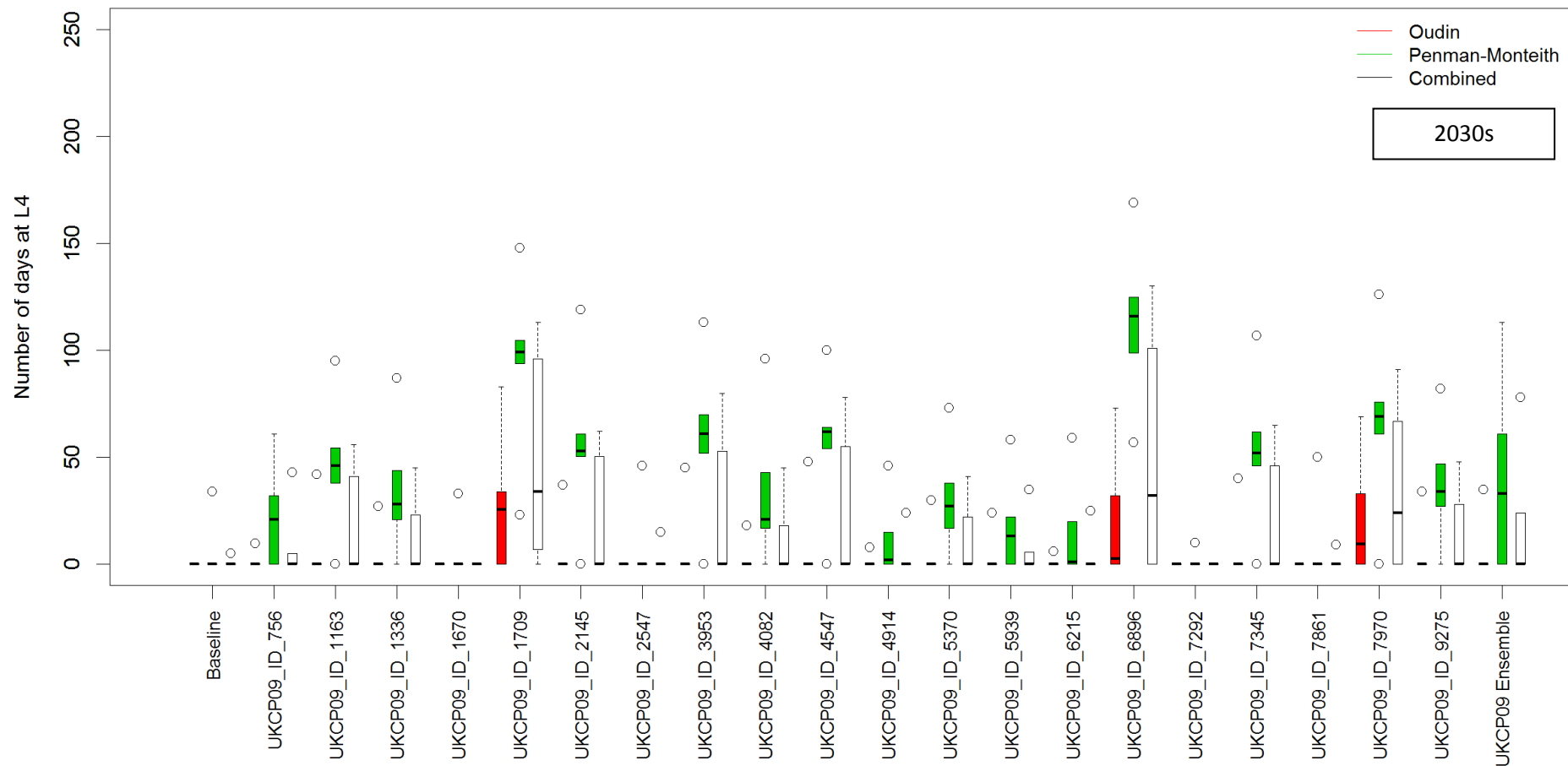
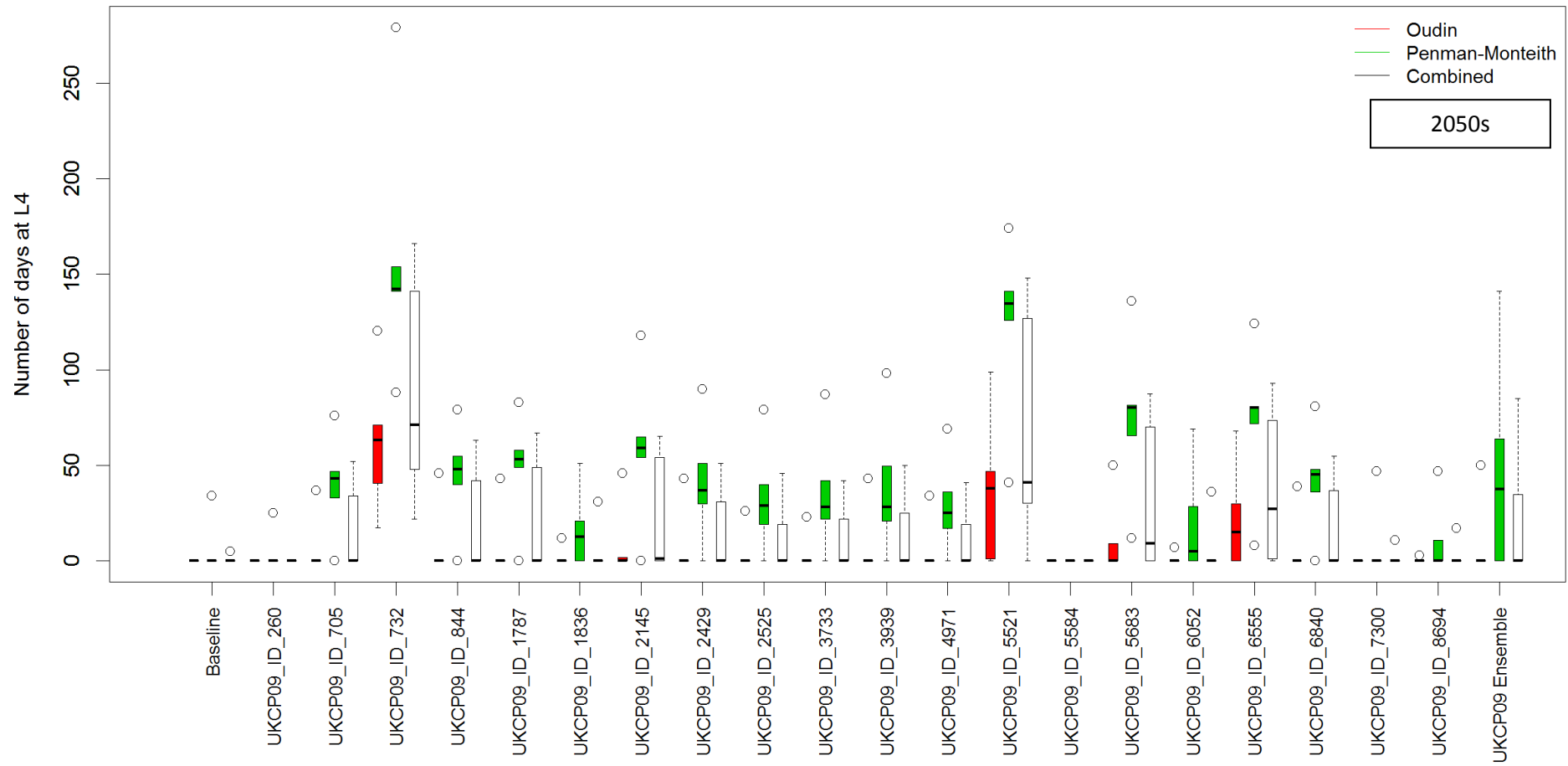
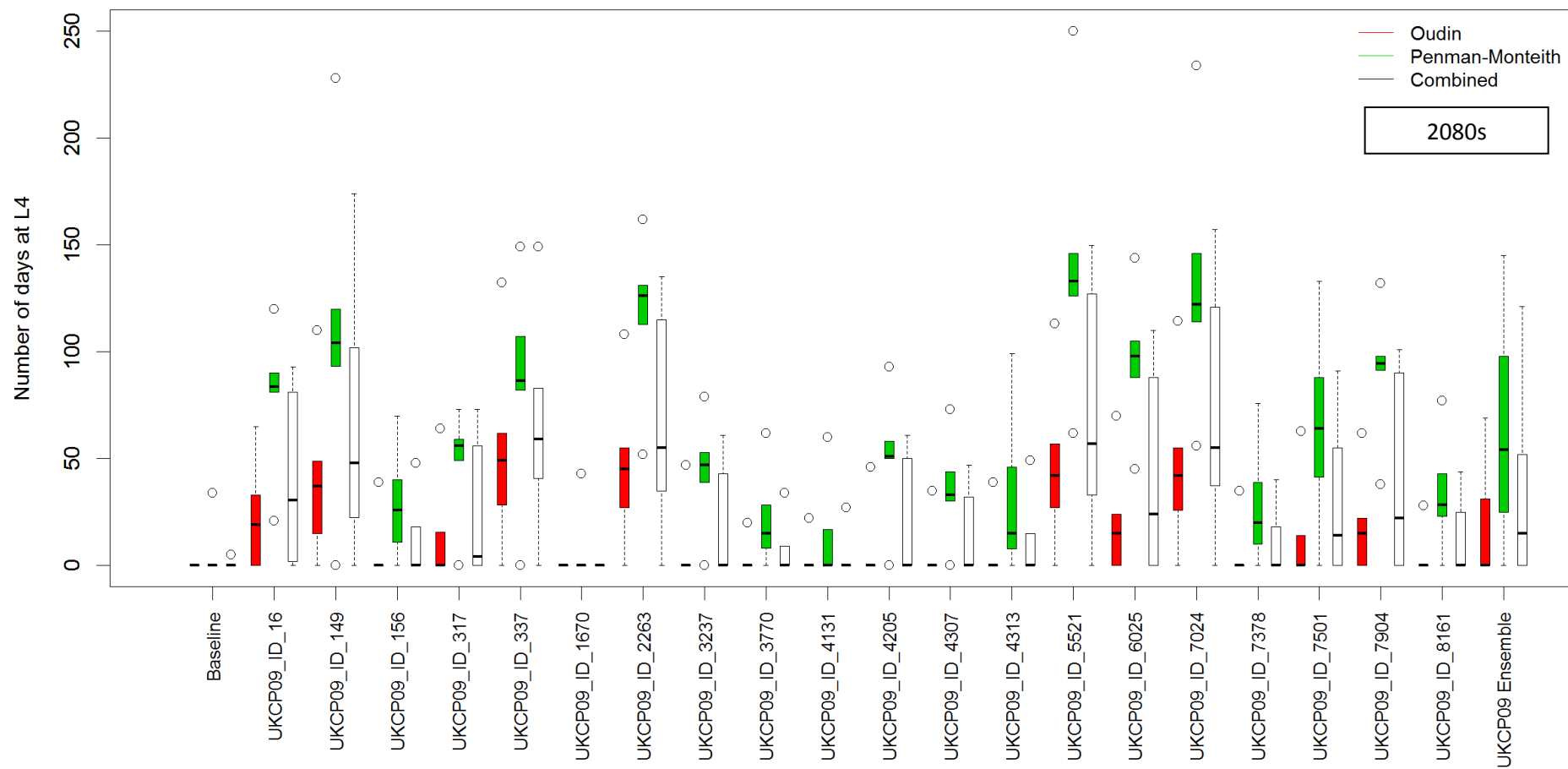


Figure 6-17: The distribution of the number of LTOACD Level 4 days (breaches) for the for each UKCP09 climate change scenario sub-sample member for the 2030s. The distribution for the baseline climatology and combined 20-member ensemble are presented at the far left and far right of each plot respectively.



**Figure 6-18: The distribution of the number of LTOACD Level 4 days (breaches) for the for each UKCP09 climate change scenario sub-sample member for the 2050s. The distribution for the baseline climatology and combined 20-member ensemble are presented at the far left and far right of each plot respectively.**



**Figure 6-19: The distribution of the number of LTOACD Level 4 days (breaches) for the for each UKCP09 climate change scenario sub-sample member for the 2080s. The distribution for the baseline climatology and combined 20-member ensemble are presented at the far left and far right of each plot respectively.**

### 6.5.2 Projected impacts of climate change on Deployable Output

Of particular interest to water resources planners is the degree to which the available resource (the Deployable Output, DO) could be impacted by climate change. In their latest (draft) Water Resource Management Plan, Thames Water (2018) reported that the estimated cost (Net Present Value, NPV) to provide additional resource of around 150 MI/d from 2030 to increase drought resilience would be of the order of £0.373bn, equivalent to ~£2.5m per MI/d. It is worth noting that as the magnitude of the additional resource required increases the associated cost per MI/d also increase because, typically, the lowest cost options are implemented first.

This DO is an input in to the supply-demand balance (see section 6.3.3) used in WRMPs to quantify the level of mitigation that may be required to maintain existing system performance. For each of the sub-sampled climate change scenarios, the demand applied in the water resources model was systematically varied in increments of 50 MI/d ( $\sim 0.6 \text{ m}^3/\text{s}$ ) for each hydrological model ensemble member until the following performance metrics were satisfied during each model simulation:

- Annual frequency of L4 breaches for hydrological model ensemble member  $\leq$  Baseline frequency of L4 breaches for hydrological model ensemble member.
- Annual frequency of L3 or L4 breaches for hydrological model ensemble member  $\leq$  Baseline frequency of either L3 or L4 breaches for hydrological model ensemble member.
- Total number of days below L4 for hydrological model ensemble member  $\leq$  Baseline total number of days below L4 for hydrological model ensemble member.
- Total number of days below L3 for hydrological model ensemble member  $\leq$  Baseline total number of days below L3 for hydrological model ensemble member.

50 MI/d ( $\sim 0.6 \text{ m}^3/\text{s}$ ) increments of demand (as a measure of Deployable Output) are equivalent to increments in Deployable Output (DO) of approximately 2.5% (2,144 MI/d,  $\sim 24.8 \text{ m}^3/\text{s}$ , is the DO reported in Thames Water's WRMP) and considered to be of sufficient resolution to reflect the relative impacts from different sources of uncertainty without increasing the computational effort significantly.

This metric provides an estimate, in the units of DO used by the UK water industry, of how the available resource is effectively reduced by climate change to maintain the current service reliability (i.e. the *frequency* and *duration* of LTOACD Level 3 and 4 transgressions – see Section 6.3.4.4). This in turn can provide an estimate as to the magnitude of additional interventions (new supply-side resources or demand management measures) that would be needed to maintain

current levels of reliability under each scenario. For each climate change scenario, the resulting thresholds of demand at which the above conditions were satisfied were combined using the conditional probabilities as described in Section 6.4.3, with the results plotted in Figure 6-20, Figure 6-21 and Figure 6-22 for the 2030s, 2050s and 2080s respectively. Figure 6-23 presents the distribution of these thresholds for all time-horizons and a summary of the results, along with the influence of the different sources of uncertainty, for each time-horizon is provided in Table 6-8.

As would be expected from the results presented in Chapter 5, and earlier in Section 6.5, the results in Figure 6-23 demonstrate that the London water resource zone may be significantly impacted by climate change. In terms of a trajectory of impacts over time, the UKCP09 projections indicate increasing reductions through to the 2080s, relative to a 1961 to 1990 baseline. As an example, the central estimates for the DO for the 2030s, 2050s and 2080s are 2,000 MI/d ( $\sim 23.1\text{m}^3/\text{s}$ ), 1,900 MI/d ( $\sim 22.0\text{m}^3/\text{s}$ ) and 1,850 MI/d ( $\sim 21.4\text{m}^3/\text{s}$ ) respectively, compared to the baseline (1961 to 1990) DO of 2,144 MI/d ( $\sim 24.8\text{m}^3/\text{s}$ ). This is equivalent to reductions of around 7%, 11% and 14%. Equating this to an economic cost, based on figures presented in Thames Water's (2018) latest draft plan highlighted above, the cost to mitigate such losses for the 2030s, 2050s and 2080s could be of the order of £360m, £610m and £735m respectively.

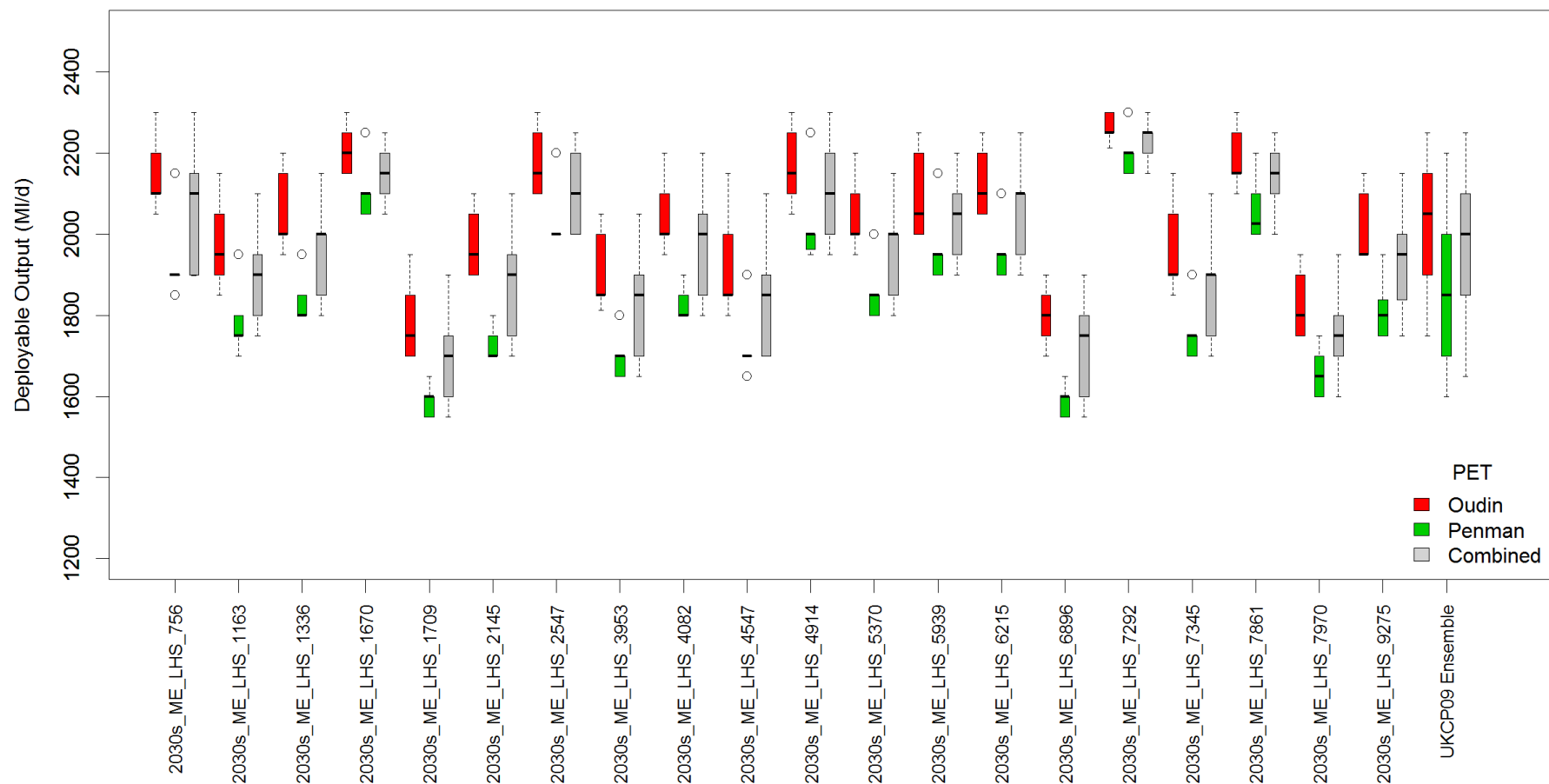
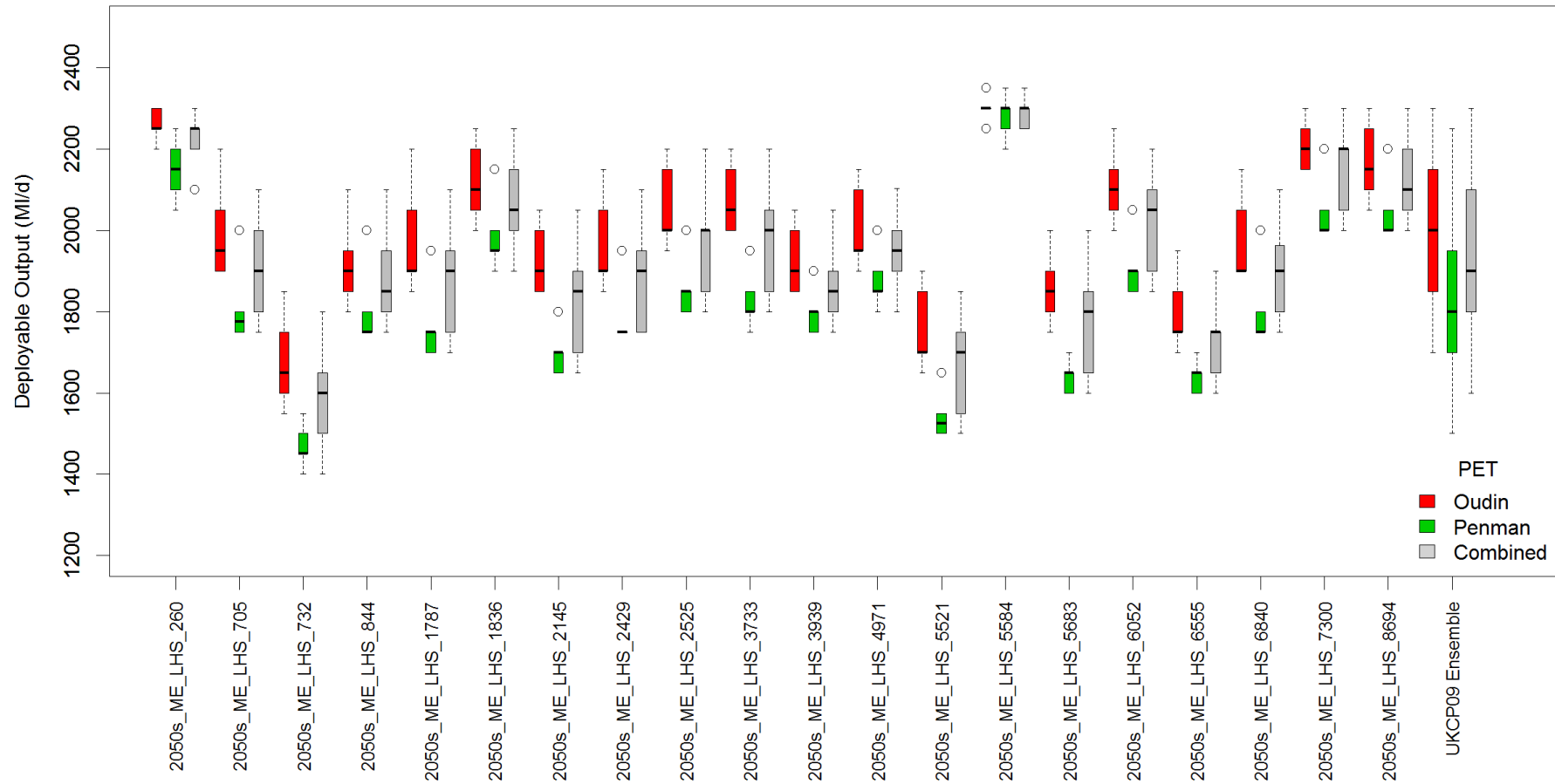


Figure 6-20: The distribution of the DO for each UKCP09 climate change scenario sub-sample member for the 2030s. Calculated as the DO at which system performance is at least as good as for the baseline condition. The distribution for the combined 20-member ensembles is presented at the far left and far right of each plot respectively. Baseline DO is 2,144 MI/d ( $\sim 24.8 \text{ m}^3/\text{s}$ ).



**Figure 6-21: The distribution of the DO for each UKCP09 climate change scenario sub-sample member for the 2050s. Calculated as the DO at which system performance is at least as good as for the baseline condition. The distribution for the combined 20-member ensembles are presented at the far left and far right of each plot respectively. Baseline DO is 2,144 MI/d ( $\sim 24.8 \text{ m}^3/\text{s}$ ).**

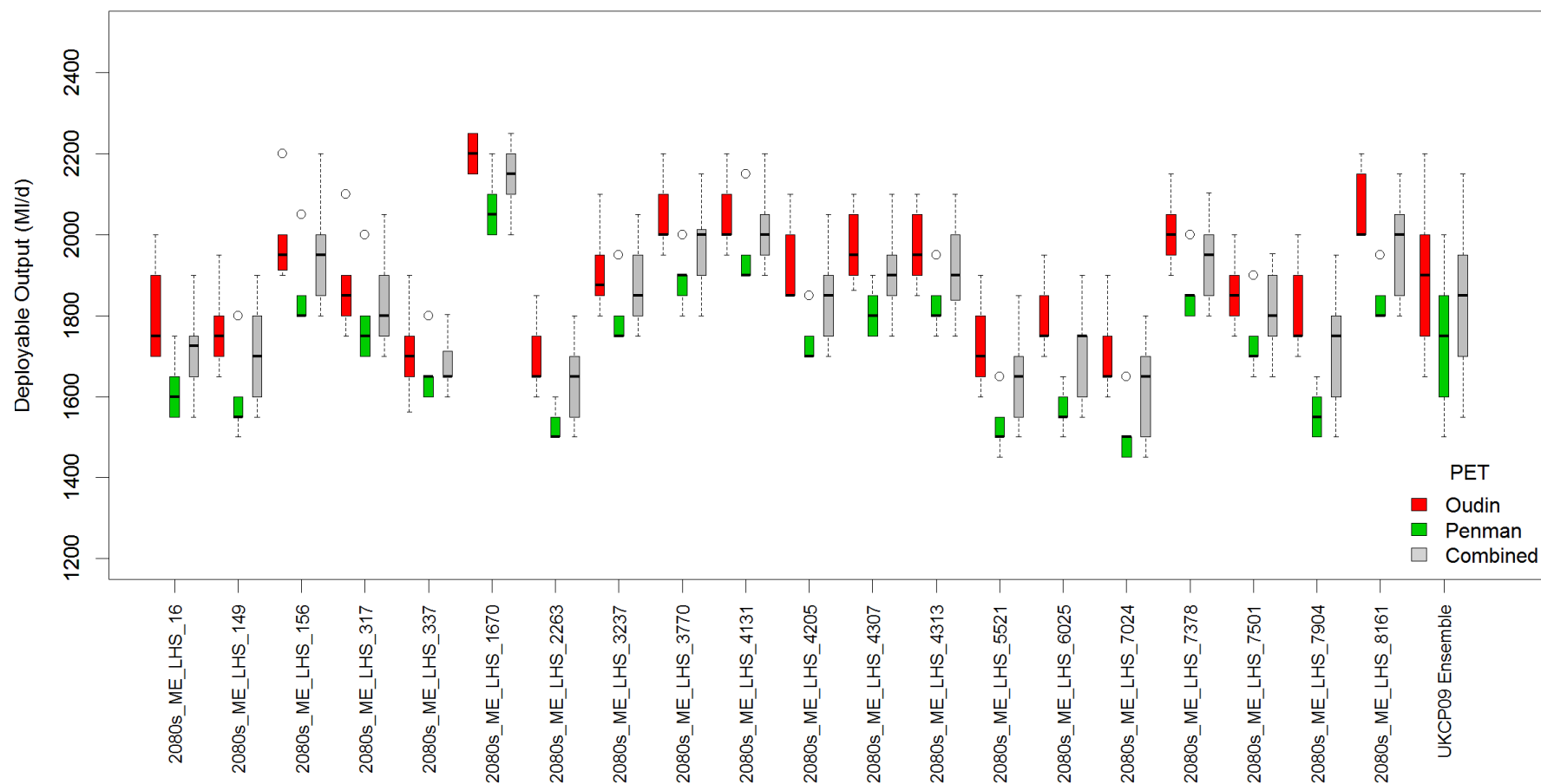
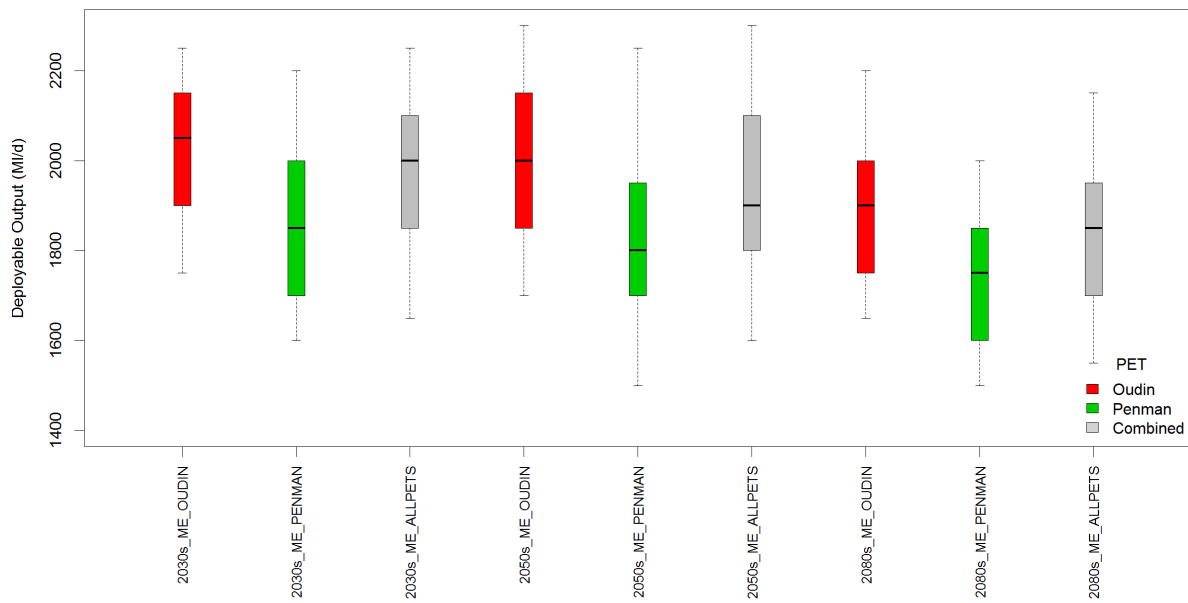


Figure 6-22: The distribution of the DO for each UKCP09 climate change scenario sub-sample member for the 2080s. Calculated as the DO at which system performance is at least as good as for the baseline condition. The distribution for the combined 20-member ensembles is presented at the far left and far right of each plot respectively. Baseline DO is 2,144 MI/d ( $\sim 24.8 \text{ m}^3/\text{s}$ ).





**Figure 6-23: The distribution of the DO for each UKCP09 sub-sample ensemble for the 2030s, 2050s and 2080s. Calculated as the DO at which system performance is at least as good as for the baseline condition. Baseline DO is 2,144 MI/d (~24.8m<sup>3</sup>/s).**

**Table 6-8: Summary of influence on Deployable Output impacts by time-horizon of different sources of uncertainty. Baseline DO is 2,144 MI/d (~24.8 m3/s).**

Impacts on Deployable Output (%) – climate change impact relative to a 1961 to 1990 baseline			2030s	2050s	2080s
<b>Combined uncertainty and influence of PET method</b>	95 <sup>th</sup> percentile estimate for the least severe climate change scenario	Both PET methods	+7%	+10%	+5%
	Central estimate across all hydrological models and climate change scenarios	Both PET methods	-7%	-11%	-14%
		Oudin PET	-5%	-7%	-12%
		mod. Penman- Mont. PET	-14%	-16%	-19%
	5 <sup>th</sup> percentile estimate for the most severe climate change scenario	Both PET methods	-28%	-35%	-32%
<b>Uncertainty due to climate scenario</b>	Climate change scenario with lowest central impact (across all hydrological model parameters)	Both PET methods	+5% (scenario 7292)	+7% (scenario 5584)	0% (scenario 1670)
	Mean of the central estimates (across all hydrological model parameter sets) of each climate change scenario		-8%	-9%	-14%
	Climate change scenario with highest central impact (across all hydrological model parameters)		-21% (scenario 1709)	-25% (scenario 732)	-23% (4 scenarios)
<b>Hydrological model parameter uncertainty</b>	Scenario with lowest sensitivity to hydrological model parameter sets (range of impacts, 5 <sup>th</sup> to 95 <sup>th</sup> percentile)	Both PET methods	Range of 7%	Range of 5%	Range of 9%
	Mean sensitivity to hydrological model parameter sets (range of impacts, 5 <sup>th</sup> to 95 <sup>th</sup> percentile)		Range of 16%	Range of 16%	Range of 16%
	Scenario with greatest sensitivity to hydrological model parameter sets (range of impacts, 5 <sup>th</sup> to 95 <sup>th</sup> percentile)		Range of 19%	Range of 19%	Range of 21%

The research presented in this chapter highlights the potentially significant influence that the method used to estimate PET may have on the predicted impacts of climate change on water resource system performance. For the 2030s, the central estimated value of DO using only the simulations using the Oudin (2005) temperature based method (MSME-200-LEE-OU simulations) is 2,050 MI/d ( $\sim 23.7 \text{ m}^3/\text{s}$ ) – a reduction of around 100 MI/d ( $\sim 1.2 \text{ m}^3/\text{s}$ ), noting that as demand increments of 50 MI/d,  $\sim 0.6 \text{ m}^3/\text{s}$ , have been used, the exact value is between 94 and 144 MI/d,  $\sim 1.1$  to  $1.7 \text{ m}^3/\text{s}$ ). Using the approximate intervention costs discussed above, this equates to a cost of between £235m and £360m.

The corresponding estimate for the MSME-200-LEE-PM simulations (using the modified Penman-Monteith method) is 1,850 MI/d ( $\sim 21.4 \text{ m}^3/\text{s}$ ), a further difference of 200 MI/d ( $\sim 2.3 \text{ m}^3/\text{s}$ ) or a further 9% reduction, and an associated additional cost implication of around £500m. For the 2050s, the estimated DOs are 2,000 MI/d ( $\sim 23.1 \text{ m}^3/\text{s}$ , Oudin) and 1,800 MI/d ( $\sim 20.8 \text{ m}^3/\text{s}$ , modified Penman – Monteith) and for the 2080s, 1,900 MI/d ( $\sim 22.0 \text{ m}^3/\text{s}$ , Oudin) and 1,750 MI/d ( $\sim 20.3 \text{ m}^3/\text{s}$ , modified Penman – Monteith) demonstrating this difference persists through the different time-horizons.

In Thames Water’s current (final) WRMP (2014), it is estimated that the climate change impacts on DO by the 2030s will be 73 MI/d ( $\sim 0.8 \text{ m}^3/\text{s}$ ), around half of the central estimate reported in this chapter. In their WARMS modelling, to reduce computational effort, Thames Water considers climate change by applying monthly flow factors to the historical river flows. These flow factors are derived by comparing historical and perturbed river flows, using a lumped CatchMod model structure to Teddington, and calculating the percentage change in the mean flow for each calendar month over the 1961 to 1990 period. The climate inputs into the WARMS CatchMod model are based on using only the Oudin method for estimating changes in PET. In the research presented here using the PDM-AQ model structure, and considering only the Oudin method for estimating PET (i.e. MSME-200-LEE-OU simulations), the reduction in DO is around 94 MI/d ( $\sim 1.1 \text{ m}^3/\text{s}$ ) and is broadly consistent with that reported by Thames Water despite the methodological differences.

When examining individual ensemble members within each epoch’s sub-sample, as in Figure 6-22, the significant variation in sub-sample ensemble members is evident. For the 2030s, the central estimate of DO (considering all MSME-200-LEE simulations) varies from 1,700 MI/d ( $\sim 19.7 \text{ m}^3/\text{s}$ , climate sub-sample member 1709) to 2,250 MI/d ( $\sim 26.0 \text{ m}^3/\text{s}$ , climate sub-sample member 7292) a range of 550 MI/d ( $\sim 6.4 \text{ m}^3/\text{s}$ ). By the 2080s this variation is from 1,650 MI/d ( $\sim 19.1 \text{ m}^3/\text{s}$ , scenarios 337, 2263, 5521, 7024) to 2,150 MI/d ( $\sim 24.9 \text{ m}^3/\text{s}$ , scenario 1670), a range of 500 MI/d

( $\sim 5.8\text{m}^3/\text{s}$ ). This analysis highlighted that 4 out of the 20 scenarios estimated the lowest DO, 1,650 MI/d ( $\sim 19.1\text{m}^3/\text{s}$ ), for the 2080s.

The influence of hydrological model uncertainty can also be seen in the distribution of estimated DO for each climate sub-sample ensemble member. In terms of the range between the 5% and 95% for the 2030s (for each climate change scenario), this varies from 150 MI/d ( $\sim 1.7\text{m}^3/\text{s}$ ) to 400 MI/d ( $\sim 4.6\text{m}^3/\text{s}$ ) in the 2030s. These are calculated by weighting the reduction in DO needed to maintain each system simulation's baseline Level of Service (the frequency and number of days below LTOACD L3 and L4 transgressions – see Section 6.3.4.4) combined for each climate scenario using the weights calculated for each system simulation as described in Section 6.4.3. By the 2080s, these ranges vary from 200 MI/d ( $\sim 2.3\text{m}^3/\text{s}$ ) to 450 MI/d ( $\sim 5.2\text{m}^3/\text{s}$ ), a difference of 250 MI/d ( $\sim 2.9\text{m}^3/\text{s}$ ). These results, and the distributions for each climate sub-sample member presented in Figure 6-22, suggest that hydrological modelling uncertainty may be enhanced or dampened by whichever climate change scenario is being considered.

These results emphasise that all of the uncertainties considered in this analysis are significant and have the potential to have a large impact on the performance of London's water resource system. By the 2080s, the overall uncertainty is very large with the estimated DO to maintain existing system performance potentially varying from 1,450 MI/d ( $\sim 16.8\text{m}^3/\text{s}$ , 5<sup>th</sup> percentile estimate for climate sub-sample member 7024) to 2,250 MI/d ( $\sim 26.0\text{m}^3/\text{s}$ , 95<sup>th</sup> percentile estimate for climate sub-sample member 1670), changes of -32% and 5% from the current estimated DO respectively. Note, as described in Section 6.3.2, the DO for London, as is standard practice in the industry, is calculated as being equal to the demand that can be met before specified LoS are breached (see Section 6.3.4.4).

These results indicate that the uncertainty associated with the climate change scenarios are larger than that associated with the aspects of hydrological modelling considered here, noting the simplifications adopted in this study including the representation of groundwater resources. However, hydrological modelling uncertainty, and the method used to estimate PET also warrant attention when planning future water supplies and yet are typically not considered by the UK water industry. These results highlight the significant challenge in planning under such large uncertainties and in particular, deciding the level and timing of investments to mitigate emerging risks. As an example, an impact that might be considered reasonably likely by the 2080s (e.g. central estimate DO impact of 14% using both PET methods) is also reasonably likely by the 2030s (e.g. central estimate DO impact of 14% using only the modified Penman-Monteith PET method).

To place these findings in the context of other potentially significant uncertainties associated with the water resources planning process, the published evidence as to the anticipated growth in demand for London is reviewed and compared to these supply-side impacts in the following section.

### *6.5.3 Uncertainty associated with the projections of the demand for water*

In their latest Water Resources Management Plan (Thames Water, 2014), Thames Water estimated that the demand for water in London could increase by 8.1% for 2035 (from a 2014 baseline). This compares to a central estimate of a 7% decrease in DO for the 2030s due to climate change reported in this chapter (noting that these impacts of climate change are relative to a 1961 to 1990 baseline).

When looking to the 2050s and 2080s, HR Wallingford Limited (2015) made use of the Office of National Statistics' (ONS, 2014) principal projection of future population change, along with their low and high fertility storylines (ONS, 2013a; ONS, 2013b), to develop projections of water demand across the UK including for the London water resource zone.

The low and high fertility storylines were requested by the Adaptation Sub-Committee<sup>3</sup> to reflect plausible upper and lower bound population estimates (HR Wallingford Limited, 2015) although it is important to stress that this is likely to reflect conservative estimates of the upper and lower bounds as fertility is only one factor that will influence future population growth with other factors such as projected mortality rates and migration. Hence, the ONS stress that these projections should not be interpreted as forecasts but rather plausible futures of population growth.

For the 2050s and 2080s, these principal, low and high fertility storylines are equivalent to the annual growth rates presented in Table 6-9. Details of the methodology used to develop these projections, including how the local authority and national scale ONS projections were translated to the water resource zone scale, is provided in HR Wallingford Limited (2015).

---

<sup>3</sup> A panel of experts from the fields of climate change, science and economics, established under the Climate Change Act 2008, to advise the Committee on Climate Change (CCC). The CCC is the independent advisor to Parliament, and the Government, on climate change issues and was also established under the Climate Change Act, 2008)

**Table 6-9: Percentage increases in population in London for the 2050s and 2080s (HR Wallingford Limited, 2015).**

<b>Population forecast</b>	<b>Population growth by 2050s</b>	<b>Population growth by 2080s</b>
Principal population (ONS, 2014)	42.5%	65.1%
Low fertility scenario (ONS, 2013a)	34.6%	46.4%
High fertility storyline (ONS, 2013b)	50.7%	86.0%

HR Wallingford Limited (2015) translated these population projections into projections of water demand using the trajectories to 2040s reported by Thames Water in their latest WRMP (2014) along with the assumptions listed in Table 6-10 for the three main components that make-up demand. The resulting increases in demand growth are presented in Table 6-11.

**Table 6-10: Household and non-household consumption and leakage assumptions used to develop projections of demand for the CCRA2 water availability report (HR Wallingford Limited, 2015).**

<b>Household demand</b>	<b>Non-household (NHH) demand</b>	<b>Leakage</b>
Metered and unmetered Per Capita Consumption (PCC) to remain at the levels as at the end of the current water resources planning horizon. All new properties are considered to be metered and adopt the PCC values reported by Thames Water for such properties.	Relationship between population growth and NHH demand growth to the 2030s developed from the latest WRMP and then projected forward to the 2050s and 2080s for each population scenario.	Per capita leakage rates remain at the levels implied for new properties during the resource planning horizon (i.e. so that the rate of change in leakage with population follows the same trend as implied by the trajectory through the resource plans).

**Table 6-11: Percentage increases in the demand for public water supplies in London for the 2050s and 2080s (HR Wallingford Limited, 2015) based on population forecasts and the assumptions detailed in Table 6-10 with regards to per capita consumption, non-household demand and leakage.**

<b>Population forecast</b>	<b>Demand growth by 2050s</b>	<b>Demand growth by 2080s</b>
Principal population	18.5% (~0.43% per annum)	29.8% (~0.37% per annum)
Low fertility scenario	14.6% (~0.34% per annum)	20.5% (~0.27% per annum)
High fertility storyline	22.6% (~0.51% per annum)	40.2% (~0.48% per annum)

As a contrast, in their recent water resources study for London, Borgomeo et al. (2014) considered annual demand growth rates of 0.70% and 1.00% to the 2050s. For their modelling of London's water resources, Walsh et al. (2015) used population estimates from the Greater London Authority's strategic plan for London for up to 2031 and then extrapolated at the same average annual growth rate of 51,000 to provide an estimate of population for 2050. This was complemented with estimates of employment growth estimates and demand per capita was altered to reflect technological advances such as improved water efficiency measures although details of these were not published.

In terms of climate change and its potential to impact the demand for water, HR Wallingford Limited (2015) assumed that demand is insensitive to climate change based on the evidence presented by studies such as Herrington et al. (1996) and UKWIR (2012). These studies both suggested there is limited evidence as to a significant weather – demand relationship. A similar assumption was also made by Borgomeo et al. (2014) for similar reasons. Walsh et al. (2015) made no reference to this issue in their publication.

Parker (2014) did identify a non-linear response of likelihood and volume of external water-use to average air temperatures that suggests climate change could have an impact on the demand for water. Parker (2014) reported an abrupt increase in the likelihood of external water-use on days above ~15°C and suggested that, as an example, by the 2080s, under a hotter/drier climate, average unmetered households could be 8% more likely to use external-water and expend around 9 litres more per day during the summer (noting that median metered household total was reported as 311 litres per day).

For the 2050s and 2080s, these estimates of demand growth demonstrate that the demand for water is likely to place greater pressures on water resources than climate change but that both are very significant and subject to large degrees of uncertainty. Borgomeo et al. (2014) reported that up to the 2050s, both climate change and population growth will increase the frequency of water shortages in London. In their research, the effects of increasing demand (due to population growth and in the absence of further efforts to limit per capita consumption) are projected to

have a greater impact on the frequency of water shortages than climate change, although a relatively simple approach to demand growth (the 0.7% and 1.0% annual growth scenarios to 2050s detailed above) was adopted for their research. Similarly, Walsh et al. (2015) reported that population growth exhibits a greater contribution to drought risk, in terms of water shortages, than climate projections, at least to the 2050s.

## 6.6 Chapter summary

The research presented in this chapter has sought to assess the relative significance of the different sources uncertainty relating to hydrological modelling and climate change considered in chapters 4 and 5 on the predicted performance of London's water resources system.

Previous studies which have considered sources of uncertainty through to their potential impact on the performance of water resource systems (e.g. Fowler et al., 2007b; Borgomeo et al., 2014; Walsh et al., 2015) have, as far as the author is aware, taken a relatively limited look at the role that different components of hydrological modelling uncertainty (including model parameterisation, performance criteria and the method used to estimate PET) may play. This is despite a large library (e.g. Wilby & Harris, 2006; Fowler et al., 2008; Kingston et al., 2009; Cloke et al., 2010; Christerson et al., 2012; Prudhomme & Williamson, 2013; Charlton & Arnell, 2014) of previous research focussed on the impacts of hydrological modelling uncertainty on river flows suggesting that their potential influence on water resources may be significant, particular when considering the very low flow sequences that drive water resources planning in the UK.

Consequently the research presented in this chapter has specifically sought to address this knowledge gap, considering the interplay between hydrological modelling and climate change uncertainty and contrasting their influence with the demand-side pressures that also face water resources planners. The time-horizons considered in this research have also extended beyond the typical 25-year planning horizons considered by the UK water industry and also beyond those by the previous studies that have looked at water resource system performance (e.g. Fowler et al., 2007b; Borgomeo et al., 2014; Walsh et al., 2015), looking at the 2030s, 2050s and 2080s.

The methodology adopted for this research has used the hydrological model and climate change ensembles from chapters 4 and 5 to quantify the performance of the water resources system of London using a simplified water resources system model. The approach taken to this research has also deliberately been framed in the language of the UK water resources planner with performance considered in terms of Levels of Service and reductions in the available resource (Deployable Output) to provide directly evidence relevant to the UK water resources industry.



### *The projected impacts for London's resource availability*

These results emphasise that all of the uncertainties considered in this analysis are significant and have the potential to have a large impact on the performance of London's water resource system. By the 2080s, the overall uncertainty is very large with the estimated DO to maintain existing system performance varying from 1,450 MI/d ( $\sim 16.8\text{m}^3/\text{s}$ , 5<sup>th</sup> percentile estimate for climate sub-sample member 7024) to 2,250 MI/d ( $\sim 26.0\text{m}^3/\text{s}$ , 95<sup>th</sup> percentile estimate for climate sub-sample member 1670), changes of -32% and 5% from the current estimated DO respectively. The relative role played by each uncertainty source varies through time and there is evidence that these sources can combine to either enhance or dampen impacts depending upon their characteristics.

### *The influence of climate change*

Chapter 5 highlighted that the UKCP09 projections exhibit a tendency towards hotter, drier summers and wetter winters in the future along with a potential delay in autumn/winter precipitation to replenish soil moisture stores and recharge aquifers. The research presented in this chapter has established that climate change is likely to have an increasing impact on the resource available to supply London, and that the uncertainty associated with these impacts is large. The central estimates of the percentage reduction in resource for the 2030s, 2050s and 2080s, according to the UKCP09 projections under a Medium Emission scenario, are of the order of 7%, 11% and 14% respectively. In terms of the range of projections, the central estimates for the 2080s range from -23% to 0%.

### *Influence of hydrological model parameter uncertainty and the PET method*

Using the Multi-Site Model Ensemble (MSME) developed in Section 4.3 has provided a means of comparing and considering the relative sensitivity of system performance to both hydrological model parameter uncertainty and uncertainty due to the method used to estimate PET. However, these estimates of sensitivities need to be treated with caution as they are conditional on the assumptions made as to the relative skill of each member of the hydrological model ensemble.

The research presented here further emphasises the potentially significant role the hydrological model uncertainty may have on water resources system performance. The influence of hydrological model uncertainty is emphasised by the fact that the distribution (between the 5<sup>th</sup> and 95<sup>th</sup> percentiles) of estimated DO for each climate sub-sample ensemble member for the 2030s varies from 150 MI/d ( $\sim 1.7\text{m}^3/\text{s}$ ) to 400 MI/d ( $\sim 4.6\text{m}^3/\text{s}$ ). By the 2080s, these ranges are from 200 MI/d ( $\sim 2.3\text{m}^3/\text{s}$ ) to 450 MI/d ( $\sim 5.2\text{m}^3/\text{s}$ ), a difference of 250 MI/d ( $\sim 2.9\text{m}^3/\text{s}$ ). The variance across climate ensemble members highlights that hydrological modelling uncertainty can be enhanced or dampened by whichever climate change scenario is being considered.

With regards to the method used to estimate PET, the central estimate of resource availability for the 2030s is around 9% higher for simulations using the Oudin (2005) temperature based PET method compared to simulations using the modified Penman-Monteith PET method (Prudhomme & Williamson, 2013). Similar differences are observed for the 2050s and 2080s. Whilst such differences may not be statistically significant they are of practical importance given the influence on the investment costs that might be associated with adaption. A difference of 9%, or 200 MI/d, could translate to an additional adaptation cost of ~£500m (NPV).

*The relative influence of supply-side and demand-side pressures and uncertainties*

For the 2030s, Thames Water's current (final) WRMP (2014) reports that the demand for water in London will increase by 8.1% for 2035 (from a 2014 baseline), compared to a central estimate of a 7% decrease in DO due to climate change (relative to a 1961 to 1990 baseline) presented in this research. When looking further ahead to the 2050s and 2080s, these estimates of demand growth demonstrate that the demand for water is likely to place greater pressures on water resources than climate change but that both are very significant and subject to large degrees of uncertainty. A recently completed study (HR Wallingford Limited, 2015) estimates the principal (ONS, 2014) projected growth in demand for the 2050s (relative to a 2014 baseline) to be around 19%, increasing to 30% by the 2080s. This compares to reductions in supplies of 11% for the 2050s and 14% for the 2080s (relative to a 1961 to 1990 baseline) estimated by the research presented in this chapter.

The uncertainty around projected population growth and the resulting impacts on demand are highlighted by the ranges reported by HR Wallingford Limited (2015), with upper and lower bound estimates for the 2080s (relative to a 2014 baseline) of 21% and 40% respectively. As for climate change, water resources planners face significant challenges in dealing with such an uncertain future and deciding the level and timing of investments to mitigate emerging risks. What might be considered as being a reasonably likely requirement to maintain adequate supplies in the 2080s might actually be needed significantly sooner and several studies are exploring more complex approaches to planning under such uncertainties and addressing these challenges (e.g. Matrosov et al., 2013; Korteling et al., 2013; Borgomeo et al., 2014). What the results in this chapter highlight, however, that alongside developing such approaches, the uncertainties associated with hydrological modelling and estimating PET, along with climate change, must not be ignored and that significant work is required to understand and quantify these uncertainties to underpin water resources planning in the future.



## **Chapter 7 Conclusions, limitations and recommendations for further research**

### **7.1 Introduction**

This research has explored the uncertainty associated with a number of the components integral to how the UK water resources industry plans for the future, moving beyond current industry practices to consider the use of alternative hydrological model structures and hydrological model parameter sets, the use of multiple performance criteria across multiple locations and time-periods and two methods for estimating PET. Further, the climate change projections (and their implicit representation of future uncertainties) used by UK water resources planners as part of their latest WRMPs have been examined. The potential influence of these different sources of uncertainty have been assessed using decision-relevant metrics, based around modelling of the London water resource system, and has examined these factors beyond the 25 year planning horizon typically adopted within the industry. These supply-side uncertainties have been compared to the uncertainty related to population growth forecasts and their projected impacts on the future demand for water.

This thesis investigates a number of knowledge gaps associated with previous academic research into water resources planning under hydrological modelling and climate change uncertainty. The majority of previous research has focussed on the impacts of uncertainty sources on river flows (e.g. New et al., 2007b; Manning et al. 2009; Cloke et al., 2010; Christerson et al., 2012; Prudhomme et al., 2012; Charlton & Arnell, 2014;) and have rarely taken their analysis as far as quantifying the impacts upon more decision-relevant metrics such as the Levels of Service and the relative frequency of triggering the drought management options that affect stakeholders. Where such studies have quantified impacts in this way only a limited number of uncertainty sources have been included (e.g. Fowler et al., 2007b; Lopez et al., 2009; Borgomeo et al.; 2014; Walsh et al. 2015) with the focus of such studies typically on other aspects of the water resources planning process such as the use of alternative decision-making methods.

A number of interesting conclusions of relevance to the UK water resources planning process have been drawn which are detailed in the following section. Alongside these, a range of limitations have been exposed, both with regards to current industry practice and some that are directly related to the methodological approach adopted in this research. These are also described in this chapter along with suggestions for future research and the additional evidence that is needed by the industry to enable improved water resources planning under uncertainty.

## 7.2 Summary of research conclusions

This section highlights the research contributions made by this thesis with regards to water resources planning in the UK.

### **Hydrological model structures, parameter uncertainty and model ensembles**

Section 4.3 described the development of a hydrological model ensemble (MSME) to generate spatially coherent river flow predictions at multiple sites of interest for water resources planning. This ensemble development adopted a Generalised Likelihood Uncertainty Estimation (GLUE) approach (Beven & Binley, 1992 - Section 4.3.1) to enable the potential, relative significance of different sources of hydrological modelling uncertainty, identified as being potentially significant in the literature review (see Section 2.3.3), to be assessed. The ensemble explored the use of two alternative hydrological model structures (Section 3.3), multiple parameter sets (Section 4.3.1.3) and two methods for estimating PET (Section 3.4.2.1). Multiple performance criteria (Section 4.3.1.1) have been used to evaluate model performance at multiple locations (Section 3.2) over multiple time-periods (Section 3.4.1). This capability was developed as part of a computationally efficient modelling framework (Section 3.3) to demonstrate the practicality of running large ensembles of simulations to provide uncertainty evaluations as part of water resource impact studies.

The modelling presented in Chapter 4 demonstrated that the hydrological response is sensitive to the choice of hydrological model structure. For the Thames catchment the PDM-AQ (Moore, 2007) model structure is shown in Section 4.3.1.4 to be more appropriate than the Thames Catchment Model (Wilby et al., 1994) based on the GLUE rejectionist approach described in Section 4.3.1.2. However, the Thames Catchment Model has been the most commonly used model structure on the Thames basin in previous water resources planning and research (e.g. Thames Water, 2013; Wilby & Harris, 2006; Environment Agency, 2009; Manning et al., 2009; Borgomeo et al., 2014; Walsh et al., 2015). This conclusion emphasises the benefits that exploring multiple model structures may offer, providing an opportunity to compare and contrast simulated responses and the evidence by which to justify the final selection of model structures for use in prediction and planning.

In evaluating hydrological modelling uncertainty the aim was to develop a model ensemble that reasonably reflected our uncertainty in the hydrological model parameters and assess its influence on decision-relevant metrics. In this analysis of the Thames parameter uncertainty was shown to be significant (Section 4.4) with the range of impacts, particularly with regards to low flows and droughts (see Sections 4.4.1 and 4.4.2), varying significantly across the ensemble of

parameter sets. For Q95 (often used as an indicator of low flows and widely used to characterise hydrological regimes for water management purposes - Charlton & Arnell, 2014) at Kingston over the 1961 to 2008 period, the MSME prediction varied from 19.9m<sup>3</sup>/s (5th percentile) to 26.8m<sup>3</sup>/s (95th percentile) with a central estimate of 23.4m<sup>3</sup>/s (compared to 23.9 m<sup>3</sup>/s from the historical record). Whilst the conditional probabilities associated with the model ensemble predictions are likely to be influenced by the performance criteria and thresholds used (discussed later in this chapter), this research does emphasise the significant sensitivity of predictions of resource availability to parameter uncertainty. The implications of hydrological model uncertainty for water resources planning is shown in Section 6.5.2 with, for example, the MSME distribution (between the 5<sup>th</sup> and 95<sup>th</sup> percentiles) of estimated DO impact due to individual climate scenarios (for the 2080s) varying from 200 MI/d (~2.3m<sup>3</sup>/s, 9% of baseline DO) to 450 MI/d (~5.2m<sup>3</sup>/s, 21% of baseline DO).

Chapter 4 included a review of the ensemble's skill in reproducing a number of water-resources relevant flow characteristics. The ensemble was shown to reproduce a range of flow characteristics well. This included periods of low flows that trigger drought management operations (see Sections 4.4.1.4, 4.4.2.1 and 4.4.2.2) across different sites and the reproduction of relatively small peak flows from short duration rainfall events when soil moisture deficits are high. However, the model ensemble struggled to reproduce wetting-up periods after sustained dry periods which was particularly evident during the autumn of 1976 – a feature also noted by Crooks & Kay's (2015) in their modelling of the Thames using the CLASSIC (a gridded derivative of the PDM model) model. Further, the ensemble demonstrated a consistent underestimation of flows during the 2001 to 2008 period. Finally, it appeared that the catchment ensemble had a tendency to 'release' more water to the river during, and just after, periods of high soil moisture deficits than was evident in the historical flow record.

It is speculated by the author that the physical processes in play at times of prolonged drought might be quite different to processes during other times and that the models used here were not able to fully reflect these changing dynamics. These processes might include river reaches experiencing significant seepage to the underlying aquifer as well as deep fissuring and cracking, especially of clay soils, providing new, faster, enhanced penetration of rainfall into the subsurface storage after a prolonged dry and hot period. In reviewing available research on drought termination (Parry et al., 2016b) highlighted that the extent to which both climate and hydrological models are able to simulate observed drought termination events remains an open question. This contextual [water resources] understanding of a model ensemble's behaviour is

important to the water resources planner in its use in prediction (e.g. estimating future resource availability under a range of future climates). For example, in this research, this tendency to simulate a recovery in river flows earlier than realised in reality has been shown to result in under-estimating the water resource implications of droughts (see the reservoir drawdown in Section 6.4.1).

The use of model ensembles may help highlight systematic limitations associated with a specific model structure, such as the challenges highlighted here with the ‘wetting-up’ processes, that could lead to the use of an alternative model structure or guide model structure refinements in the future. However, for this thesis, it should also be borne in mind that the difficulties in reproducing such aspects of the flow regime may not be solely related to the models but may also be related to other uncertainties such as those associated with model inputs (e.g. weather data and the historical flow record itself - e.g. see McMillan et al. 2012; Beven & Binley, 2013; Westerberg et al., 2016). Some of the related uncertainties (e.g. influence of abstractions and discharges, uncertainty in flow measurements) may, proportionally, have a greater impact when river flows are at their lowest. The use of multiple time-periods as part of the model evaluation process may help in diagnosing changes in catchment response (e.g. systematic changes in abstractions and discharges within the catchment) or issues associated with flow measuring stations that have not been considered in this thesis.

### **Climate change projections for water resources planning**

Climate variability and climate change are implicitly intertwined making it difficult to separate one from the other. The latest WRMPs produced by water companies in England and Wales used two alternative sources of climate projections for water resources planning: the eleven-member Future Flows (Prudhomme et al., 2012) transient climate ensemble and the UKCP09 (Murphy et al., 2009b) ‘probabilistic’ projections. The UKCP09 projections do not allow for the explicit quantification of climate variability and effectively superimpose a climate change signal upon the historical climate record. In contrast, the Future Flows projections are transient, bias-corrected (for the purposes of hydrological modelling, Prudhomme et al., 2012), continuous time-series based on the outputs from regional climate models and are considered to implicitly include both climate change and natural climate variability. Both of these sources of climate change evidence have been examined in this research through applying a “top-down” approach to climate change impacts assessment (see section 2.3.2.5).

#### *Overview of the projected changes in climate for the Thames basin*

The Future Flows and UKCP09 climate projections show a tendency towards hotter, drier

summers and wetter winters in the future along with a potential delay in autumn/winter precipitation to replenish soil moisture stores and recharge aquifers. The UKCP09 ensemble estimate changes in summer (June to August) precipitation, by the 2080s under the Medium Emission scenario, to vary from a reduction of 65% to an increase of 37%, with changes in summer temperature ranging from +1 to +9 degrees Celsius. The Future Flows projections estimate mean precipitation reductions to range from a reduction of 63% to an increase of 47% for specific summer months by the 2080s. These large ranges highlight the very high uncertainty associated with the climate projections currently used by water resources planners in the UK.

#### *Overview of the projected changes in river flows for the Thames basin*

The Thames catchment is anticipated to be significantly impacted by such changes in climate with all five locations studied showing significant decreases in low flows. Across the five sites the central estimate of reduction in Q95 by the 2030s, using the UKCP09 probabilistic projections, is between 6.1% and 23.2% (with a 13.2% reduction at Kingston). In terms of a trajectory of impacts over time the UKCP09 projections lead to increasing reductions on low to medium flows through to the 2080s, with the reduction in Q95 across the five sites varying from 11.5% to 34.1% (with a 21.3% reduction at Kingston) by this time horizon. Consistent with the findings of Charlton & Arnell (2014) flashier parts of the basin are shown to be particularly vulnerable to large reductions in summer flows benefitting less from the baseflow contributions provided by the storage of winter precipitation. In terms of drought events, the MSF at Kingston for the worst event in the 1961 to 2008 period reduces from 21.1 m<sup>3</sup>/s (central estimate from the model ensemble under the historical climate data) to 16.1 m<sup>3</sup>/s (central estimate from the model ensemble under the 2080s perturbed historical data), a reduction of 5 m<sup>3</sup>/s. Alongside the direct impacts of reduced water availability, lower minimum flows imply less volume for dilution and, hence, potentially higher pollution concentrations downstream of point discharges such as wastewater treatment works (Whitehead et al., 2009).



This part of the research (see Section 5.4.2, Figure 5-14) also highlighted the relative uncertainties associated with the climate change projections relative to the uncertainty associated with the hydrological model parameters. Using the GLUE rejectionist approach described in Section 4.3.1.2 to develop the MSME, this research suggests that for all time-horizons (2030s, 2050s, 2080s) the former is much larger (as shown by the variation between climate scenarios) but that hydrological parameter uncertainty is still significant (i.e. reflected in the variation within each individual climate scenario). This is consistent with the evidence reported from other studies (e.g. Wilby and Harris, 2006; New et al., 2007a; Manning et al., 2009; Christerson et al., 2012; Charlton & Arnell, 2014).

#### *Use of the Future Flows transient climate projections for water resources planning*

The transient projections of precipitation and PET contained within the Future Flows ensemble (and the resulting flows generated using these projections) have been compared with the equivalent series from the historical record (the 1961 to 2008 period). This comparison demonstrated a general tendency for the Future Flows ensemble members to be wetter than the historical record in the Thames basin. This tendency was in terms of both the mean annual precipitation and more critically for water resources planners, lower mean maximum deficits recorded during drier periods than in the historical record (Section 5.3.2). Further, this research has demonstrated significant variation between ensemble members in their estimates of PET for which the Future Flows project (Prudhomme et al., 2012) adopted the Penman-Monteith (Allen et al., 1998) equation (see Section 2.2.2.2). Some Future Flows ensemble members can be seen, on average, to be both significantly wetter (Table 5-1) and exhibit significantly higher evapotranspiration rates (Table 5-2) than the historical record.

Whilst inspection of the flow duration curves and changes in mean summer flows for low flow years between 1961 and 2008 (see Section 5.4.1.2) highlighted variability across ensemble members, none of the Future Flows ensemble members contain dry periods that are as severe as the historical record. Overall, whilst the most severe events in these transient sequences might be of interest to 'stress' test a system, significant caution should be exercised before using the Future Flows climatology for water resources planning in the Thames basin (such as undertaken by Huskova et al., 2016) because the reproduction of persistent dry periods may be limited. These findings reinforce concerns (e.g. Goodess, 2013; Rocheta et al., 2014; Watts et al., 2015) that the outputs from the underlying climate models that were used to develop the UKCP09 and the Future Flows projections have limitations in reproducing the meteorological blocking patterns that are significant in the formation of extended droughts.

### **Estimating potential evapotranspiration (PET)**

The uncertainty associated with estimating PET has been largely overlooked by previous research that has considered impacts beyond river flows onto the resulting implications on maintaining public water supplies. This research has applied two contrasting PET methods, a modified Penman-Monteith approach (Prudhomme & Williamson, 2013) and an empirical temperature based method, the Oudin (2005) method. Due to its complete reliance on temperature the Oudin formula could be considered less appropriate under more extreme future climates than methods such as the modified Penman-Monteith which use multiple weather inputs. However, that could also be considered an advantage in that it does not require the input of a range of weather variables such as relative humidity that may be less reliably predicted by climate models (Prudhomme & Williamson, 2013).

Both PET methods suggested a general tendency (trend) for increasing rates of PET through the 1961 to 2008 period (Section 3.4.2.1) consistent with the general warming across the UK reported by others (e.g. Jenkins et al., 2009). No obvious trend is evident within the precipitation record (Section 3.4.2.2). Typically, summer PET values calculated using the Oudin method peak at higher monthly rates, including for the particularly dry periods such as the summers of 1965, 1991, 1992, 1997, 2003 and 2005. However, the opposite is true during the hot, dry summer of 1976 and also for the summer of 1990.

Annual estimates are higher using the modified Penman-Monteith method with the annual mean PET between 1961 to 2008 estimated to be 636mm and 590mm using the modified Penman-Monteith and Oudin method respectively. The differences are primarily related to estimates during winter months with mean October-to-April values of 193mm (modified Penman-Monteith) and 156mm (Oudin) which would affect the estimated rates of groundwater recharge (and resulting river baseflow during the following summer). Finally, the modified Penman-Monteith method was shown to exhibit greater temporal variability (Figure 3-6) with larger cumulative deficits and surpluses relative to their respective long-term average 365-day cumulative PET values than the Oudin method.

The method used for estimating PET is therefore potentially significant (as suggested by Haxton & Young, 2012), particularly during hot, dry periods, and could lead to differences in the overall water balance maintained by hydrological models of the Thames catchment. However, in this research, the model structures were able to deliver models that performed similarly under both methods with model parameterisation able to compensate for such water balance differences when reproducing the historical flow record. However, when taken forward into the climate

change modelling, the method used to estimate PET has been shown to exert a significant influence on the change in projected river flows across multiple locations in the Thames catchment. Larger reductions in the Q95 metric are generated using PET changes estimated using the modified Penman-Monteith method than using the Oudin method. Using the Oudin method the central estimate of reduction by the 2080s (considering all sampled UKCP09 climate scenarios) in Q95 at Kingston is 19.1% compared to 22.3% using the modified Penman-Monteith method. This difference due to the PET method varied across the sampled climate scenarios (Figure 5-14) with the central estimates for one scenario in the 2080s being reductions of 31.6% under the modified Penman-Monteith method and 24.7% under the Oudin method. This difference across climate scenarios reflects the relative contribution from the changes in precipitation and PET with the largest differences due to the choice of PET method evident in those scenarios with the greatest predicted increase in summer PET.

When taking these impacts through to the water resources modelling the central estimate of resource availability for the 2080s (under a Medium Emission scenario) is around 12% higher using the Oudin PET method compared to using the modified Penman-Monteith PET method. This demonstrates that further research is needed to better understand and quantify the changes in PET under a future climate as part of water resources planning. This may be particularly important for those water resource systems largely dependent upon largely permeable catchments, such as the Thames, which may be especially sensitive.

### **Water resources modelling and planning**

Chapter 6 used a simplified model of London's water resource system to quantify, using metrics directly relevant to water resource planners, the impacts on the performance of the London water resources system from the different sources of uncertainty considered in chapters 4 and 5. These impacts were then compared to the uncertainty associated with future population projections and the resulting demand for water.

The results presented in Chapter 6 (e.g. Figure 6-23) demonstrate that the London water resource zone may be significantly impacted by climate change. In terms of a trajectory of impacts over time, the water resources modelling using the UKCP09 projections indicate (central predictions) increasing reductions in Deployable Output, relative to a 1961 to 1990 baseline, of 7% (2030s), 11% (2050s) and 14% (2080s). The potential investment costs (NPV) to mitigate such reductions in supplies could be of the order of £735m, noting that the investment cost per MI/d typically increases as the scale of the additional resource needed increases.

The uncertainties in predicted resource availability are large with the estimated reduction in DO to maintain existing system performance varying from 1,450 MI/d ( $\sim 16.8 \text{ m}^3/\text{s}$ , 5th percentile estimate for an individual, relatively hot and dry, climate scenario in the 2080s) to 2,250 MI/d ( $26.0 \text{ m}^3/\text{s}$ , 95th percentile estimate for a an individual, relatively cool and wet climate scenario in the 2080s). These are equivalent to changes of -32% and 5% in respectively and a potential investment cost (NPV) difference of over £2Bn.

The relative role played by the different sources of uncertainty were shown to vary for different time-horizons. In terms of climate change uncertainty the MSME central estimate of resource reduction by the 2080s for each climate scenario varied from -23% to 0%. The influence of hydrological model uncertainty is emphasised by the distribution (between the 5<sup>th</sup> and 95<sup>th</sup> percentiles) of estimated DO for each climate sub-sample ensemble member for the 2080s. This varied from 200MI/d ( $\sim 2.3 \text{ m}^3/\text{s}$ ) to 450 MI/d ( $\sim 5.6 \text{ m}^3/\text{s}$ ), a difference of 250 MI/d ( $\sim 2.9 \text{ m}^3/\text{s}$ ). This variance across climate ensemble members highlights that hydrological modelling uncertainty can be enhanced or dampened by whichever climate change scenario is being considered, further emphasising the importance of explicitly including hydrological modelling uncertainty as part of water resources planning.

In their current (final) WRMP (2014) Thames Water estimate that the demand for water in London will increase by 8.1% by 2035 (from a 2014 baseline). This compares to a central estimate of a 7% decrease in Deployable Output for the 2030s (relative to a 1961 to 1990 baseline) due to climate change reported in this thesis. For the 2050s and 2080s (noting that Thames Water only considered a 25 year planning horizon in their last WRMP) the results from this research were compared against those from a recent study that considered the projected impacts of population growth (HR Wallingford Limited, 2015). The study by HR Wallingford Limited estimated that the principal (ONS, 2014) projected growth in demand for the 2050s to be around 19%, increasing to 30% by the 2080s (from a baseline of 2014). This compares to reductions in supplies of 11% for the 2050s and 14% for the 2080s (relative to a 1961 to 1990 baseline) estimated here.

These results highlight that the uncertainties related to both projected population growth and climate change present significant challenges, potentially of a similar magnitude, that will require significant adaptation to maintain public water reliable supplies to London in the future.

Furthermore, planning under such large uncertainties presents very significant challenges in deciding upon the level and timing of investments to mitigate emerging risks. As demonstrated in Chapter 6, what might be considered as being a reasonably likely requirement to maintain adequate supplies in the 2080s might actually be needed significantly sooner.

### 7.3 Limitations of research

The following section highlights limitations identified during the writing of this thesis including reflections as to what these might mean for current industry practice.

#### **A single case study**

This research has only considered a single river basin in detail (the Thames catchment) and London's water supply system. Each catchment and water supply system will include unique conditions. Thames Water's London water resource zone is particularly dependent upon abstractions from the River Thames which are then stored in large raw water reservoirs and the operation of the system is subject to a bespoke operating agreement regulating the flows in the lower Thames (Thames Water, 2013). Therefore, other systems, particularly those with significantly less storage, both through aquifers and surface water reservoirs, may exhibit very different responses and degrees of sensitivity to the sources of uncertainty considered in this thesis.

#### **Hydrological model structures**

In modelling the Thames catchment only two alternative hydrological model structures have been used. These model structures were chosen because they are used extensively in the UK and were considered appropriate for the Thames catchment. However, CatchMod and PDM are only two of the very large number of hydrological models that exist and other hydrological model structures may suggest different responses to those presented in this thesis and may offer more appropriate process representations to address some of the limitations identified here (e.g. reproducing the recovery after a sustained period of low flow).

The CatchMod model structure was selected because it is the model structure used by both Thames Water (2014) and the Environment Agency (2009) to model the Thames catchment and has been used in previous research investigating topics related to this thesis (e.g. Wilby & Harris, 2006; Manning et al., 2009; Borgomeo et al., 2014). However, all model realisations using this model structure were rejected under the GLUE (Beven & Binley, 1992) based approach to considering uncertainty adopted in this research (see Section 4.3.1) which suggests that its application in the Thames warrants review.

The second model structure, PDM-AQ, was selected because it is an established model structure used in a number of academic and industry studies (e.g. Moore, 2007; UKWIR, 2007; Reynard et al., 2010; Prudhomme et al., 2012). PDM-AQ provides a conceptual contrast to the CatchMod structure in its formulation, particularly the use of slow and fast pathways and a probability

distributed soil moisture store. The form of PDM adopted here was one considered appropriate for use in permeable, groundwater dominated catchments, adopting a quadratic non-linear store (as used in the TCM model structure) for the 'slow' (groundwater) pathway considered appropriate for unconfined aquifers (Moore, 2002) but challenges using this model structure (described above) have also been identified.

However, the relative performance of the model structures will be partly dependent upon the design of the GLUE based approach to considering uncertainty, including the choice of performance metrics and conditioning periods. It is worth noting that each model structure, with the four HRUs used, has 20 'free' parameters and a limitation of this research is that it has not undertaken systematic diagnostic evaluation of the model structures to better understand why certain characteristics of the hydrological signature are particularly difficult to reproduce. This could have included investigation into the sensitivity of the parameterisation schemes and individual model parameters to the conditioning data and performance metrics used.

Currently water resource planners do not take into account hydrological model parameter uncertainty and whilst this research has shown its potential influence on the water resources planning process, further research is needed to better understand what is important and why. As suggested in Chapter 4, this could lead to using a different set of behavioural models depending on the conditions (states) of the catchment to achieve improved hydrological performance across a wider range of conditions. However, it is worth highlighting that such an approach would present challenges in identifying when to switch between predictive modes and the subsequent use of the resulting flow series in water resources system modelling where temporal coherence in each simulated flow series is typically necessary.

#### **Uncertainties associated with historical/observational climate data**

Historical climate data have been taken from the Ensembles E-Obs version 3.0 (Haylock et al., 2008) and the UK Met Office (2013) 'historical' gridded datasets but the uncertainty associated with historical weather data has not been considered. For the calculation of PET two contrasting methods have been used; the Oudin (2005) temperature based method and the modified Penman-Monteith equation (Prudhomme & Williamson, 2013). These two methods were used because they can exploit the variables available in the UKCP09 climate change projections, albeit assuming no change in mean wind speed for the modified Penman-Monteith method. However, several other methods (e.g. see Prudhomme & Williamson, 2013) exist that were not considered here.

During the course of this research CEH has made available a new gridded precipitation and PET datasets available (e.g. GEAR - Keller et al., 2015; CHESS – Robinson et al., 2016) but these have not been used in the impacts modelling. These new data could provide a useful comparison of the data used here albeit that these are likely to be underpinned by similar weather station data to those used in the E-Obs datasets. For PET, Robinson et al. (2016) made use of the Penman-Monteith equation (Allen et al., 1998) noting that a modified form of this equation has been used in this thesis.

It is also worth emphasising that for water resource planners it is AET rather than PET that influences water availability. Transpiration and evaporation often occur at less than their potential rate due to crop stress arising from limited soil moisture availability (e.g. Rushton et al., 2006; Zhao et al., 2013) which will be the case during the extended droughts that are of particular interest to water resource planners. Consequently, further work examining how hydrological models reproduce these effects and estimate AET from PET during such periods would be of value to planners, particularly when considering the use of these models for simulating climate change scenarios that include significant degrees of warming.

#### **Other sources of uncertainty not considered in this thesis**

As well as the limited consideration of the sources of uncertainty discussed earlier, there are other uncertainties that have not been considered. These include the uncertainty associated with the historical flow records, emission uncertainty associated with the climate change projections, the uncertainty due to future land use changes and behavioural changes in the use of water and future changes in environmental flow requirements.

#### **Simplifications of London's water resources system**

A number of simplifications to the representation of London's water resources infrastructure have been made. Consequently, whilst the modelling is appropriate for meeting the objectives of this research to assess and compare the relative influence of different sources of uncertainty, the absolute values of the supply-demand balance should not necessarily be considered to reflect all the operational constraints (e.g. pump operations, pipe capacities) that are taken into account in the more detailed WARMS analysis undertaken by Thames Water (Thames Water, 2014).

As part of these simplifications the groundwater available to support London's water supplies was assumed to provide resource that was equal to the dry year deployable output. This is the maximum rate at which groundwater sources are estimated (Thames Water, 2013) to be available during dry periods. More detailed analysis of groundwater resource, particularly under the more

severe climate change scenarios, would identify whether more or less groundwater resource (when used conjunctively with the surface water sources) would be available during droughts.

### **Limitations of available climate change projections**

The adequacy of the climate models that underpin the UKCP09 and Future Flows climate projections, and the methods used to apply them in this research, have significant limitations. It is recognised that the Global Climate Models (GCMs) that underpin the UKCP09 and Future Flows projections do not include important earth system processes (e.g. climate-induced emissions from wetlands and permafrost) and have weaknesses in their representation of aspects, such as atmospheric blocking (Scaife et al., 2010, Sillmann et al., 2017), of particular importance to water resource planners. Such limitations are likely to remain an issue in the near term given that the CMIP5 GCMs that will underpin the next set of climate projections for the UK (UKCP18), continue to underestimate the observed blocking frequency over Europe (Anstey et al., 2013). Further, the coarse scale of the GCMs means that they cannot provide information as to the spatial structure of temperature and precipitation in areas of complex topography and land use distribution, and their depiction of regional and local atmospheric circulations, and representation of processes at high frequency temporal scales, are such that they are not considered appropriate for direct impacts modelling (Christensen et al., 2007).

The analysis of the Future Flows projections presented in Chapter 5 highlighted the resulting issues with regards to their skill in reproducing natural climate variability, including persistent dry periods, for water resources planning. Section 2.2.2.2 described the Future Flows project's use of bias-correction (Newton et al., 2012) of the UKCP09 RCM outputs to develop climate projections considered appropriate for hydrological modelling (Prudhomme et al., 2012). This need for bias-correction emphasises the limitations in the modelling of physical processes within the underlying RCMs and is itself a further source of uncertainty (e.g. challenges in accounting for skewness and kurtosis in the bias-correction process as well as its sensitivity to the choice of calibration period - Teutschbien & Siebert, 2013; Lafon et al., 2013). The use of such approaches also assumes that biases relative to historical observations will be constant into the future under a changing climate (e.g. Teutschbein & Seibert, 2013).

The UKCP09 projections through the perturbation of historical rainfall and PET do result in the system being tested against droughts more intense than those in the baseline record. Using a perturbation approach is commonly adopted primarily because of its simplicity and that, in principle, it combines one-step downscaling and bias correction (Willems et al., 2012), assuming that relative changes from climate models are more reliable than absolute values. However, the



use of the UKCP09 change factors makes the fundamental assumption that change in extreme drought events will be similar in magnitude to the changes in mean climate on which the UKCP09 projections have been conditioned. Furthermore, future variability is effectively considered unaffected with the sequence and frequency of weather patterns unchanged.

This thesis emphasises that more research is needed to enable water resource planners to capture and plan for significant changes in hydro-climatic characteristics relevant to water security assessments, particularly the changes in drought events. Recent academic research (e.g. Turner et al., 2014; Borgomeo et al., 2015) and industry studies (Environment Agency, 2015) have demonstrated the use of synthetic drought sequences to test the sensitivity of water resource systems to a range of droughts of different characteristics to those within the historical record. Borgomeo et al. (2015) and Turner et al. (2014) have both demonstrated how uncertain information from climate models can be coupled with this understanding, as part of a “bottom-up” approach (see section 2.3.2.5) to climate change, to compare decision alternatives based on their robustness to changing drought characteristics. However, such approaches are still limited by the evidence implicit in the climate projections, particularly the weaknesses in the underlying climate models as to the generation of droughts (e.g. Goodess, 2013; Rocheta et al., 2014; Watts et al., 2015). Whilst the current water resources planning methods for adapting to climate change may be considered to make the most out of the limited sources of evidence currently available, the lack of evidence as to the potential impacts of climate change on drought formation and persistence remains a critical knowledge gap for water resources planners.

### **Approach to quantifying uncertainty**

The approach to evaluating hydrological model performance and model rejection used here has been explicitly described and is easily reproducible but may be considered to incorporate a degree of subjectivity. Six performance metrics have been considered over multiple catchment locations and multiple time-periods under a Generalised Likelihood Uncertainty Estimation (GLUE) methodology (Beven & Binley, 1992) where for a model to be selected it had to pass pre-defined thresholds for all criteria.

Limitations with this approach include that the six performance metrics may result in an aspect of double counting in that they may reinforce similar aspects of model behaviour. However, even where independent measures can be identified the relative weights applied to these would incorporate some aspect of subjectivity - it is difficult, a priori, to understand the relative importance of different measures on the final objectives of the modelling (in this case the projected supply – demand balance for London). Furthermore, whilst the thresholds for not

‘rejecting’ a model, partly informed by previous studies, were set in advance of the evaluation process, these may also be considered subjective (Pianosi & Wagener, 2016). The approach also assumes that hydrological modelling uncertainty (i.e. which models are considered behavioural and their relative contribution in prediction) remains constant into the future under different flow conditions. Further, other sources of uncertainty (e.g. use of multiple climate models and alternative methods of bias-correction, the uncertainty associated with the water resources system modelling and the uncertainty associated with historical input data) have not been considered. Therefore, the results presented in this thesis may be better considered as reflecting the sensitivity of model results to the different sources of uncertainty rather than as absolute estimates of actual uncertainties.

#### 7.4 Suggestions for future research

##### **Hydrological model structures, modelling uncertainty and estimating PET**

These research findings encourage further examination of the appropriate use of alternative hydrological model structures and the accuracy of the available data for the Thames to understand issues such as why the model ensemble struggled to reproduce well the wetting-up after the 1976 drought and the consistent underestimation of flows during the 2001 and 2008 period. More generally, evidence as to the strengths and weaknesses of alternative model structures and methods for estimating PET is required to enable improvement of the representation of key processes, particularly during droughts.

These issues are becoming increasingly critical as the industry is moving towards the use of risk-based methods adopting synthetic and stochastic weather sequences for understanding water resource system resilience (Southern Water, 2014; UKWIR, 2016c) rather than simply relying upon the historical flow record. This is placing even greater reliance on the quantification of uncertainties and the adequacy of hydrological models, in particular their performance during conditions which may be significantly different to which their predictive ‘skill’ has been evaluated. These advanced planning methods therefore will require improvements in the representation of critical hydrological processes such as groundwater – surface water interactions and the treatment of artificial influences (abstractions and discharges) with a view to improved representation of droughts and their termination. Further, the combination of considering droughts more severe than the historical record and the increasing temperatures suggested by climate projections highlight the need for further research around estimating PET along with how hydrological models calculate AET from PET during periods of reduced soil moisture. This may be

especially important for water resource systems, such as London, that are supplied by largely permeable catchments.

As the industry starts to consider the increased sharing of water and reform of water abstraction policy to better manage available resources into the future (DEFRA, 2014), spatially distributed models may offer benefits beyond simply improved model performance. Spatially distributed models could support the provision of spatially coherent flows at all locations within a large catchment, better spatial representation of all abstractions and discharges and enable the identification of particular river reaches or sub-catchments at particular risk from over abstraction (e.g. DEFRA, 2014). The application of GLUE could be explored further to enable identification of spatially distributed models considered capable of providing coherent outputs at a variety of locations and to identify models that could be applied at a variety of spatial decision scales. Consequently, it is recommended that further work is invested in the application of semi-distributed models, particularly in heavily modified catchments such as the Thames, to support better water resources planning in the future.

### **Climate change projections**

Significant research and development (e.g. IPCC, 2013) continues into the climate models used to generate scenarios of climate change. For water resources planners, models considered sufficiently skilful in generating the full range of droughts that could occur under a changed climate remain the ultimate goal that would enable a step-change in the planning of robust and resilient water resources systems. In the interim it is critical that the limitations of the climate change ‘products’ currently available to planners are clearly identified and communicated. Further, such projections should be interpreted in the context of the specific vulnerabilities by water resource systems, as suggested by Turner et al. (2014) and Borgomeo et al. (2015).

The next evolution of UK climate change projections is due in 2018 and it is recommended that detailed examination as to their appropriateness for water resources planning for different parts of the UK is undertaken. The limitations of these projections, and any derived outputs, must be clearly communicated to the industry and stakeholders. The examination should focus on the projections’ skill in relation to multi-season droughts and place this evidence in the context of different types of systems and their vulnerability to droughts. The research presented in this thesis starts to provide an approach for providing water resources relevant evidence to support the use of climate change projections as part of the water resources planning process.

## **Quantification of uncertainty**

This research has examined a range of uncertainties associated with estimating future resource availability and provides a mechanism for identifying key vulnerabilities and uncertainties in the planning of future water resources. However, the limitations of the approach to characterising and quantifying uncertainty adopted in this research are recognised by the author and other approaches, such as the use of formal Bayesian statistics (e.g. Kavetski and Fenicia, 2011; Euser et al., 2013) or the extended GLUE approach (e.g. Liu et al., 2009) may offer alternative methodologies. These alternative approaches are subject to their own limitations and typically require additional evidence not always readily available to the water resources planner. However, it is recommended that further research explores the alternative methods of quantifying uncertainty, particularly in heavily modified catchments of interest to water resources planning, to enable improved quantification of uncertainty as part of the water resources planning process.

## **7.5 Concluding remarks**

Uncertainty in future conditions and changing stakeholder priorities are encouraging the development of new methods for decision making in water resources planning (Matrosov et al., 2013; Korteling et al., 2013; Huskova et al., 2016; UKWIR, 2016b; UKWIR 2016c). Several decision-making methods are being explored to deal with planning under deep or severe uncertainty (Walker et al., 2013) associated with future conditions and delivering solutions that are considered to favour robustness (perform well across a range of plausible futures) rather than precise optimality for one (or a few) specific future condition. Such decision-making approaches require defensible estimates of the uncertainty associated with all of the inputs into the decision-making process. These include the supply-side uncertainties that have been investigated in this thesis along with the uncertainty related to other factors such as the demand for future water, future energy costs and the costs of securing capital investment.

This thesis demonstrates the influential role that different sources of supply-side uncertainty might play in the water resources planning process, including the uncertainties associated with hydrological modelling and estimating PET that are largely ignored by the UK water resources industry at this time. It is recognised that whatever decision-making approaches are adopted within the industry, the treatment of uncertainty cannot be completely exhaustive. Therefore, the quantification of the different sources of uncertainty needs to be balanced with pragmatic approaches that can be readily applied by the UK water resources industry. As the industry moves towards the use of these extended decision-making methods it is hoped that this research

contributes valuable evidence as to which sources of uncertainty warrant particular scrutiny and how such uncertainties might be assessed.

## References

- Abbot B., M. & Refsgaard C., J. (1996) Distributed hydrological modelling, Kluwer Academic.
- Abbott, M. B., Bathurst, J. C., Cunge, J. A., OConnell, P. E. and Rasmussen, J. (1986) An introduction to the European Hydrological System -- Systeme Hydrologique Europeen, "SHE", 1: History and philosophy of a physically-based, distributed modelling system, *Journal of Hydrology*, 87(1–2), pp. 45–59.
- Allen, R.G. & Pruitt, W.O. (1991) FAO-24 reference evapotranspiration factors. *Journal of Irrigation and Drainage Engineering*, 117(5), pp.758-773.
- Allen, R.G., Pereira, L.S. , Raes, D. and Smith, M. (1998) Crop evapotranspiration - Guidelines for computing crop water requirements, Rome: FAO.
- Amarasinghe, P., Liu, A., Egodawatta, P., Barnes, P., McGree, J. and Goonetilleke, A. (2016) Quantitative assessment of resilience of a water supply system under rainfall reduction due to climate change. *Journal of Hydrology* 540 pp. 1043-1052.
- Anglian Water Services (2014). Water resources Management Plan 2014.
- Anstey, J.A., Davini, P., Gray, L.J., Woollings, T.J., Butchart, N., Cagnazzo, C., Christiansen, B., Hardiman, S.C., Osprey, S.M. and Yang, S. (2013) Multi-model analysis of Northern Hemisphere winter blocking: Model biases and the role of resolution. *Journal of Geophysical Research: Atmospheres*, 118(10), pp.3956-3971.
- Arnell, N.W. and Reynard, N.S. (1996) The effects of climate change due to global warming on river flows in Great Britain. *Journal of Hydrology*. 397-424. 183.
- Arnell, N.W., Hudson, D.A. and Jones, R.G. (2003) Climate change scenarios from a regional climate model: Estimating change in runoff in southern Africa. *Journal of Geophysical Research: Atmospheres*, 108(D16).
- Arnell, N.W. (2011) Incorporating Climate Change Into Water Resources Planning in England and Wales1. *JAWRA Journal of the American Water Resources Association*, 47(3), pp.541–549.
- Arnold, J. G., Srinivasan, R., Muttiah, R. S. and Williams, J. R. (1998) Large area hydrologic modeling and assessment part 1: model development, *Journal of the American Water Resources Association*, 34(1).

- Arnold, S., Attinger, S., Frank, K. and Hildebrandt, A. (2009) Uncertainty in parameterisation and model structure affect simulation results in coupled ecohydrological models, *Hydrology and Earth System Sciences*, 13(10), pp. 1789–1807.
- Bahremand, A., De Smedt, F., Corluy, J., Liu, Y., Poorova, J., Velcicka, L. and Kunikova, E. (2007) WetSpa Model Application for Assessing Reforestation Impacts on Floods in Margecany-Hornad Watershed, Slovakia, *Water Resources Management*, 21(8), pp. 1373–1391.
- Bárdossy, A., Huang, Y. & Wagener, T. (2016) Simultaneous calibration of hydrological models in geographical space. *Hydrology and Earth System Sciences*, 20(7), pp.2913–2928.
- Bastola, S., Murphy, C. & Sweeney, J. (2011) Evaluation of the transferability of hydrological model parameters for simulations under changed climatic conditions. *Hydrology and Earth System Sciences Discussions*, 8(3), pp.5891–5915.
- Beguiría, S., Vicente-Serrano, S. M., Reig, F. and Latorre, B. (2014) Standardized precipitation evapotranspiration index (SPEI) revisited: parameter fitting, evapotranspiration models, tools, datasets and drought monitoring, *International Journal of Climatology*. John Wiley & Sons, Ltd, 34(10), pp. 3001–3023. d
- Bell, V.A. & Moore, R.J. (1998) A grid-based distributed flood forecasting model for use with weather radar data: Part 1. Formulation. *Hydrology and Earth System Sciences* (2/3), p.265.
- Bell, V. A., Kay, A. L., Jones, R. G. and Moore, R. J. (2007a) Development of a high resolution grid-based river flow model for use with regional climate model output, *Hydrology and Earth System Sciences*, 11(1).
- Bell, V. A., Kay, A. L., Jones, R. G. and Moore, R. J. (2007b) Use of a grid-based hydrological model and regional climate model outputs to assess changing flood risk, *International Journal of Climatology*, 27(12).
- Bell, V.A., Kay, A.L., Jones, R.G., Moore, R.J. and Reynard, N.S. (2009) Use of soil data in a grid-based hydrological model to estimate spatial variation in changing flood risk across the UK. *Journal of Hydrology*, 377(3), pp.335-350.
- Bell, V.A., Gedney, N., Kay, A.L., Smith, R.N., Jones, R.G. and Moore, R.J. (2011) Estimating potential evaporation from vegetated surfaces for water management impact assessments using climate model output. *Journal of Hydrometeorology*, 12(5), pp.1127-1136.

- Bergström, S. & Graham, L.P. (1998) On the scale problem in hydrological modelling. *Journal of Hydrology*, 211(1), pp.253-265.
- Beven, K.J., and Kirkby, M.J. (1979) A physically based, variable contributing area model of basin hydrology. *Hydrological Services Bulletin* 24(1): 43-69.
- Beven, K.J., Kirkby, M.J., Schofield, N. and Tagg, A.F. (1984) Testing a physically-based flood forecasting model (TOPMODEL) for three UK catchments. *Journal of Hydrology*, 69(1-4), pp.119-143.
- Beven, K., Calver, A. and Morris, E.M. (1987) The Institute of Hydrology Distributed Model (IHDM). 97. Wallingford, UK: Institute of Hydrology.
- Beven, K. & Binley, A. (1992) The future of distributed models: Model calibration and uncertainty prediction. *Hydrological Processes*, 6(3), pp.279–298.
- Beven, K. (1999) How far can we go in distributed hydrological modelling? *Hydrology and Earth System Sciences*, 5(1), pp.1–12.
- Beven, K. & Freer, J. (2001a) A dynamic TOPMODEL. *Hydrological Processes*, 15(10), pp.1993–2011.
- Beven, K. & Freer, J. (2001b) Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *Journal of Hydrology*, 249(1-4), pp.11–29.
- Beven, K.J. (2001) *Rainfall-Runoff Modelling: The Primer*, Wiley.
- Beven, K. (2006) A manifesto for the equifinality thesis. *Journal of Hydrology*, 320(1–2), pp.609–618.
- Beven, K.J., Smith, P.J. and Freer, J.E. (2008) So just why would a modeller choose to be incoherent?. *Journal of hydrology*, 354(1), pp.15-32.
- Beven, K. & Westerberg, I. (2011) On red herrings and real herrings: disinformation and information in hydrological inference. *Hydrological Processes*, 25(10), pp.1676-1680.
- Beven, K. & Binley, A. (2013) GLUE: twenty years on. *Hydrological Processes*.



- Beven, K. & Smith, P. (2014) Concepts of information content and likelihood in parameter calibration for hydrological simulation models. *Journal of Hydrologic Engineering*, 20(1), p.A4014010.
- Birks, C., Owen, M. and Arkell, B. (2001) London's water resources: threat or opportunity. *Area*, 33(1), pp.95-97.
- Blaney, H.F. & Criddle, W.D. (1950) Determining water needs from climatological data. *USDA Soil Conservation Service*. SOS--TP, USA, pp.8–9.
- Blazkova, S. & Beven, K. (2009) A limits of acceptability approach to model evaluation and uncertainty estimation in flood frequency estimation by continuous simulation: Skalka catchment, Czech Republic. *Water Resources Research*, 45(12), p.W00B16.
- Bloomfield, J.P., Allen, D.J. & Griffiths, K.J. (2009) Examining geological controls on baseflow index (BFI) using regression analysis: An illustration from the Thames Basin, UK. *Journal of Hydrology*, 373(1–2), pp.164–176.
- Bloschl, G. & Montanari, A. (2010) Climate change impacts – throwing the dice?, *Hydrological Processes*, 24, 374–381.
- Booij, M.J. (2005) Impact of climate change on river flooding assessed with different spatial model resolutions. *Journal of Hydrology*, 303(1–4).
- Boorman, D.B. & Sefton, C.E.M. (1997) Recognising the uncertainty in the quantification of the effects of climate change on hydrological response. *Climatic Change*, 35(4), pp.415–434.
- Borgomeo, E., Hall, J. W., Fung, F., Watts, G., Colquhoun, K. and Lambert, C. (2014) Risk-based water resources planning: Incorporating probabilistic nonstationary climate uncertainties, *Water Resources Research*, 50(8), pp. 6850–6873.
- Borgomeo, E., Pflug, G., Hall, J. W. and Hochrainer-Stigler, S. (2015) Assessing water resource system vulnerability to unprecedented hydrological drought using copulas to characterize drought duration and deficit, *Water Resources Research*, 51(11), pp. 8927–8948.
- Bormann, H., Breuer, L., Giertz, S., Huisman, J. A. and Viney, N. R. (2009) Spatially explicit versus lumped models in catchment hydrology - experiences from two case studies in NATO Science for Peace and Security Series -C: Environmental Security, *Uncertainties in Environmental Modelling and Consequences for Policy Making*. Croatia: Springer Netherlands, pp. 3–26.

- Boughton, W.C. (1989) A review of the USDA SCS curve number method. *Soil Research*, 27(3), pp.511–523.
- Boyle, D. P., Gupta, H. V. & Sorooshian, S. (2000) Toward improved calibration of hydrologic models: Combining the strengths of manual and automatic methods. *Water Resources Research*, 36, 3663-3674.
- Brigode, P., Oudin, L. and Perrin, C. (2013) Hydrological model parameter instability: A source of additional uncertainty in estimating the hydrological impacts of climate change?. *Journal of Hydrology*, 476, pp.410-425.
- Broderick, C., Matthews, T., Wilby, R.L., Bastola, S. and Murphy, C. (2016) Transferability of hydrological models and ensemble averaging methods between contrasting climatic periods. *Water Resources Research*, 52(10), pp.8343-8373.
- Brown, C., Ghile, Y., Lavery, M. and Li, K. (2012) Decision scaling: Linking bottom-up vulnerability analysis with climate projections in the water sector, *Water Resources Research*, 48(9).
- Cameron, D. (2006) An application of the UKCIP02 climate change scenarios to flood estimation continuous simulation for a gauged catchment in the northeast of Scotland, UK (with uncertainty). *Journal of Hydrology*, 328(1-2), 212-226.
- Carter, T.R. and Mäkinen, K. (2011) Approaches to climate change impact, adaptation and vulnerability assessment: towards a classification framework to serve decision-making. MEDIATION Technical Report No. 2.1, Finnish Environment Institute (SYKE), Helsinki, Finland, 70 pp.
- Chaney, N. W., Herman, J. D., Reed, P. M. and Wood, E. F. (2015) Flood and drought hydrologic monitoring: The role of model parameter uncertainty, *Hydrology and Earth System Sciences*, 19(7), pp. 3239–3251.
- Charlton, M.B. & Arnell, N.W. (2014) Assessing the impacts of climate change on river flows in England using the UKCP09 climate change projections. *Journal of Hydrology*.
- Chow, V.T., Maidment, D.R. & Mays, L.W. (1988) *Applied hydrology*, Singapore: McGraw-Hill.
- Christensen, J., Carter, T., Rummukainen, M. and Amanatidis, G. (2007) Evaluating the performance and utility of regional climate models: the PRUDENCE project, *Climatic Change*, 81(0), pp. 1–6.

- Christierson, B. v., Vidal, J.P. & Wade, S.D. (2012) Using UKCP09 probabilistic climate information for UK water resource planning. *Journal of Hydrology*, 424–425, pp.48–67.
- Chu, X. & Steinman, A. (2009) Event and continuous hydrologic modeling with HEC-HMS. *Journal of Irrigation and Drainage Engineering*, 135(1), pp.119-124.
- Clark, M. P., Slater, A. G., Rupp, D. E., Woods, R. A., Vrugt, J. A., Gupta, H. V, Wagener, T. and Hay, L. E. (2008) Framework for Understanding Structural Errors (FUSE): A modular framework to diagnose differences between hydrological models, *Water Resources Research*. American Geophysical Union, 44.
- Clark, M.P., Kavetski, D. & Fenicia, F. (2012) Reply to comment by K. Beven et al. on Pursuing the method of multiple working hypotheses for hydrological modeling. *Water Resources Research*, 48(11).
- Clark, M.P., Wilby, R.L., Gutmann, E.D., Vano, J.A., Gangopadhyay, S., Wood, A.W., Fowler, H.J., Prudhomme, C., Arnold, J.R. and Brekke, L.D. (2016) Characterizing uncertainty of the hydrologic impacts of climate change. *Current Climate Change Reports*, 2(2), pp.55-64.
- Cloke, H. L., Jeffers, C., Wetterhall, F., Byrne, T., Lowe, J. and Pappenberger, F. (2010) Climate impacts on river flow: projections for the Medway catchment, UK, with UKCP09 and CATCHMOD, *Hydrological Processes*. John Wiley & Sons, Ltd., pp. 3476–3489.
- Cloke, H. L., Wetterhall, F., He, Y., Freer, J. E. and Pappenberger, F. (2013) Modelling climate impact on floods with ensemble climate projections, *Quarterly Journal of the Royal Meteorological Society*, 139(671), pp. 282–297.
- Collins, M. (2007) Ensembles and probabilities: a new era in the prediction of climate change. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365(1857), pp.1957–1970.
- Coxon, G., Freer, J., Wagener, T., Odoni, N. A. and Clark, M. (2014) Diagnostic evaluation of multiple hypotheses of hydrological behaviour in a limits-of-acceptability framework for 24 UK catchments, *Hydrological Processes*, 28(25), pp. 6135–6150.
- CPDN (2016). ClimatePrediction.Net (CPDN). Available at: <http://climateprediction.net> [Accessed November 18, 2016].

- Crooks, S.M. & Kay, A.L. (2015) Simulation of river flow in the Thames over 120 years: Evidence of change in rainfall-runoff response? *Journal of Hydrology: Regional Studies*, 4(PB), pp.172–195.
- Crooks, S.M. & Naden, P.S. (2007) CLASSIC: a semi-distributed rainfall-runoff modelling system. *Hydrology and Earth System Sciences* 11(1), pp.516–531.
- Culley, S., Noble, S., Yates, A., Timbs, M., Westra, S., Maier, H.R., Giuliani, M. and Castelletti, A. (2016) A bottom-up approach to identifying the maximum operational adaptive capacity of water resource systems to a changing climate. *Water Resources Research*, 52(9), pp.6751-6768.
- Darch, G., Arkell, B. & Tradewell, J. (2011) Water resource planning under climate uncertainty in London.
- Dawson, C.W., Abrahart, R.J. and See, L.M. (2007) HydroTest: a web-based toolbox of evaluation metrics for the standardised assessment of hydrological forecasts. *Environmental Modelling & Software*, 22(7), pp.1034-1052.
- DEFRA (2014). The Impact of Water Abstraction Reform The Impact of Water Abstraction Reform.
- Deser, C., Phillips, A., Bourdette, V. and Teng, H. (2012) Uncertainty in climate change projections: the role of internal variability, *Climate Dynamics*. Springer-Verlag, 38(3–4), pp. 527–546.
- Dessai, S. and Hulme, M. (2004) Does climate adaptation policy need probabilities?. *Climate policy*, 4(2), pp.107-128.
- Dessai, S. and van der Sluijs, J. (2007) Uncertainty and Climate Change Adaptation – a Scoping Study. A report for the Netherlands Environmental Assessment Agency, in a joint project between Copernicus Institute of Utrecht University, Tyndall Centre for Climate Change Research, UK and School of Environmental Sciences, University of East Anglia, Norwich, UK.
- Diaz-Nieto, J. and Wilby, R.L. (2005) A comparison of statistical downscaling and climate change factor methods: impacts on low flows in the River Thames, United Kingdom. *Climatic Change*, 69(2-3), pp.245-268.
- Directive: 92/43/EEC. Council Directive 92/43/EEC of 21 May 1992 on the conservation of natural habitats and of wild fauna and flora.
- Directive 2000/60/EC. Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 Establishing a framework for Community action in the field of water policy, OJL 327.

Droogers, P. and Allen, R.G. (2002) Estimating reference evapotranspiration under inaccurate data conditions. *Irrigation and drainage systems*, 16(1), pp.33-45.

Duan, Q. & Qingyun, V. (1992) The shuffled complex evolution (SCE-UA) method. Department of Hydrology and Water Resources, University of Arizona Tucson.

Ekström, M., Jones, P. D., Fowler, H. J., Lenderink, G., Buishand, T. A. and Conway, D. (2007) Regional climate model data used within the SWURVE project? 1: projected changes in seasonal patterns and estimation of PET, *Hydrology and Earth System Sciences Discussions*, 11(3), pp. 1069–1083.

Environment Agency (2008). Water resources in England and Wales - current state and future pressures. Environment Agency, Bristol, UK. Available at:  
<http://webarchive.nationalarchives.gov.uk/20140329213237/http://cdn.environment-agency.gov.uk/geho1208bpas-e-e.pdf>

Environment Agency (2009). Applying probabilistic climate change information to strategic resource assessment Climate Change Science programme *Science report : SC050045*. Environment Agency, Bristol, UK. Available at:  
[https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/291618/scho0309bprh-e-e.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/291618/scho0309bprh-e-e.pdf)

Environment Agency (2012). Water Resources Planning Guideline. Environment Agency, Ofwat, Defra and the Welsh Government . Published by Environment Agency, Bristol, UK.

Environment Agency (2013). Water stressed areas – final classification, Environment Agency and Natural Resources Wales. Environment Agency, Bristol, UK. Available at:  
[https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/244333/water-stressed-classification-2013.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/244333/water-stressed-classification-2013.pdf)

Environment Agency (2015). Understanding the performance of water supply systems during mild to extreme droughts. Project ref: SC120048/S. Environment Agency, Bristol, UK. Available at:  
<https://www.gov.uk/government/publications/understanding-the-performance-of-water-supply-systems-during-mild-to-extreme-droughts>

Environment Agency (2017). Water Resources Planning Guideline: Interim update. Environment Agency, Bristol, UK. Available at: <https://naturalresources.wales/media/681612/interim-wrpg-update-final-april-2017.pdf>

- Euser, T., Winsemius, H. C., Hrachowitz, M., Fenicia, F., Uhlenbrook, S. and Savenije, H. H. G. (2013) A framework to assess the realism of model structures using hydrological signatures, *Hydrology and Earth System Sciences*, 17(5), pp. 1893–1912.
- Fowler, H. & Kilsby, C. (2007) Using regional climate model data to simulate historical and future river flows in northwest England. *Climatic Change*, 80(3), pp.337–367.
- Fowler, H.J., Blenkinsop, S. & Tebaldi, C. (2007a) Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *International Journal of Climatology*, 27(12), pp.1547–1578.
- Fowler, H.J., Kilsby, C.G. & Stunell, J. (2007b) Modelling the impacts of projected future climate change on water resources in north-west England. *Hydrology and Earth System Sciences*, 11(3), pp.1115–1124.
- Fowler, H., Tebaldi, C. & Blenkinsop, S. (2008) Probabilistic estimates of climate change impacts on flows in River Eden, Cumbria. In BHS 10th *National Hydrology Symposium*, Exeter. pp. 416–423.
- Freer, J., Beven, K. & Ambroise, B. (1996) Bayesian Estimation of Uncertainty in Runoff Prediction and the Value of Data: An Application of the GLUE Approach. *Water Resources Research*, 32(7), pp.2161–2173.
- Freer, J. E., McMillan, H., McDonnell, J. J. and Beven, K. J. (2004) Constraining dynamic TOPMODEL responses for imprecise water table information using fuzzy rule based performance measures, *Journal of Hydrology*, 291(3–4), pp. 254–277.
- Goodess, C.M. (2013) How is the frequency, location and severity of extreme events likely to change up to 2060?. *Environmental science & policy*, 27, pp.S4-S14.
- Gosling, S.N. and Arnell, N.W. (2016) A global assessment of the impact of climate change on water scarcity. *Climatic Change*, 134(3), pp.371-385.
- Green, W.H. & Ampt, G.A. (1911) Studies on Soil Physics. *The Journal of Agricultural Science*, 4(1), pp.1–24.
- Griffiths, J., Keller, V., Morris, D. And Young, A. (2008) Continuous Estimation of River Flows (CERF) - Science Report SC030240. Environment Agency.

- Gupta, H., Sorooshian, S. & Yapo, P. (1999) Status of Automatic Calibration for Hydrologic Models: Comparison with Multilevel Expert Calibration. *Journal of Hydrologic Engineering*, 4(2), pp.135–143.
- Gustard, A., Bullock, A. & Dixon, J.M. (1992) Low flow estimation in the United Kingdom, Institute of Hydrology.
- Hall, J. W., Watts, G., Keil, M., de Vial, L., Street, R., Conlan, K., OConnell, P. E., Beven, K. J. and Kilsby, C. G. (2011) Towards risk-based water resources planning in England and Wales under a changing climate, *Water and Environment Journal*. Blackwell Publishing Ltd .
- Hannaford, J. (2015) Climate-driven changes in UK river flows: A review of the evidence. *Progress in Physical Geography*, 39(1), pp.29-48.
- Haxton, T. & Young, A. (2012) Future Flows and Groundwater Levels: Development of CERF. Science Report/Project Note SC090016/PN7. Wallingford, Oxfordshire: CEH.
- Hayashi, A., Akimoto, K., Sano, F., Mori, S. and Tomoda, T. (2010) Evaluation of global warming impacts for different levels of stabilization as a step toward determination of the long-term stabilization target. *Climatic Change*, 98(1-2), p.87.
- Haylock, M. R., Hofstra, N., Klein Tank, A. M. G., Klok, E. J., Jones, P. D. and New, M. (2008) A European daily high-resolution gridded data set of surface temperature and precipitation for 1950-2006, *J.Geophys.Res. AGU*, 113, p. D20119.
- Hellebrand, H. & van den Bos, R. (2008) Investigating the use of spatial discretization of hydrological processes in conceptual rainfall runoff modelling: a case study for the meso-scale. *Hydrological Processes*, 22(16), pp.2943–2952.
- Herrington, P. (1996) *Climate Change and the Demand for Water*, London.
- Hornberger, G.M. & Spear, R.C. (1980) Eutrophication in peel inlet: I. The problem-defining behavior and a mathematical model for the phosphorus scenario. *Water Research*, 14(1), pp.29–42.
- Horton, R.E. (1941) An approach toward a physical interpretation of infiltration-capacity. *Soil Science Society of America Journal*, 5(C), pp.399–417.

Hough, M.N. & Jones, R.J.A. (1997) The United Kingdom Meteorological Office rainfall and evaporation calculation system: MORECS version 2.0-an overview. *Hydrology and Earth System Sciences Discussions*, 1(2), pp.227–239.

HR Wallingford Limited (2014). Thames Water: Water Resources Support. Independent Review of WARMS2.

HR Wallingford Limited (2015). CCRA2 : Updated projections for water availability for the UK, Available at: <https://documents.theccc.org.uk/wp-content/uploads/2015/09/CCRA-2-Updated-projections-of-water-availability-for-the-UK.pdf>.

Huskova, I., Matrosov, E. S., Harou, J. J., Kasprzyk, J. R. and Lambert, C. (2016) Screening robust water infrastructure investments and their trade-offs under global change: A London example, *Global Environmental Change*, 41, pp. 216–227.

ICA (2012) Global water security. IC-coordinated paper. Intelligence Community Assessment, Washington, DC.

IPCC (2013). Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.

Jakeman, A. J., Littlewood, I. G. & Whitehead, P. G. (1990) Computation of the instantaneous unit hydrograph and identifiable component flows with application to two small upland catchments. *Journal of Hydrology*, 117, 275-300.

Jenkins, G., Perry, M., Prior, J. and Centre, M. O. H. (2009) The climate of the UK and recent trends. Hadley Centre, Met Office, Exeter. Available at: <http://ukclimateprojections.metoffice.gov.uk/media.jsp?mediaid=87932&filetype=pdf>

Jiménez-Cisneros, B.E., Oki, T., Arnell, N.W., Benito, G., Cogley, J.G., Doll, P., Jiang, T. and Mwakalila, S.S. (2014). Freshwater resources. In: Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Field, C.B., V.R. Barros, D.J. Dokken, K.J. Mach, M.D. Mastrandrea, T.E. Bilir, M. Chatterjee, K.L. Ebi, Y.O. Estrada, R.C. Genova, B. Girma, E.S. Kissel, A.N. Levy, S. MacCracken, P.R. Mastrandrea, and L.L. White (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 229-269.



Johnson, M.S., Coon, W.F., Mehta, V.K., Steenhuis, T.S., Brooks, E.S. and Boll, J. (2003) Application of two hydrologic models with different runoff mechanisms to a hillslope dominated watershed in the north-eastern US: a comparison of HSPF and SMR. *Journal of Hydrology*, 284(1), pp.57-76.

Jones, P., Harpham, C., Kilsby, C., Glenis, V. and Burton, A. (2009) Projections of future daily climate for the UK from the Weather Generator. UK.

Kavetski, D. & Fenicia, F. (2011) Elements of a flexible approach for conceptual hydrological modeling: 2. Application and experimental insights. *Water Resources Research*, 47(11).

Kay, A.L. & Davies, H.N. (2008) Calculating potential evaporation from climate model data: A source of uncertainty for hydrological climate change impacts. *Journal of Hydrology*, 358(3–4), pp.221–239.

Kay, A. L., Davies, H. N., Bell, V. A. & Jones, R. G. (2009) Comparison of uncertainty sources for climate change impacts: flood frequency in England. *Climatic Change*, 92, 41-63.

Kay, A.L. & Jones, D.A. (2010) Transient changes in flood frequency and timing in Britain under potential projections of climate change. *International Journal of Climatology*.

Keller, V. D. J., Tanguy, M., Prosdocimi, I., Terry, J. A., Hitt, O., Cole, S. J., Fry, M., Morris, D. G. and Dixon, H. (2015) CEH-GEAR: 1 Km resolution daily and monthly areal rainfall estimates for the UK for hydrological and other applications, *Earth System Science Data*, 7(1), pp. 143–155.

Kendon, M., McCarthy, M., Jevrejeva, S. and Legg, T. (2016) State of the UK Climate 2015, Met Office, Exeter, UK.

Kingston, D. G., Todd, M. C., Taylor, R. G., Thompson, J. R. and Arnell, N. W. (2009) Uncertainty in the estimation of potential evapotranspiration under climate change, *Geophysical Research Letters*, 36(20), p. L20403.

Korteling, B., Dessai, S. & Kapelan, Z. (2013) Using Information-Gap Decision Theory for Water Resources Planning Under Severe Uncertainty. *Water Resources Management*, 27(4), pp.1149–1172.

Krueger, T., Freer, J., Quinton, J. N., Macleod, C. J. A., Bilotta, G. S., Brazier, R. E., Butler, P. and Haygarth, P. M. (2010) Ensemble evaluation of hydrological model hypotheses, *Water Resources Research*. AGU, p. W07516.

- Lafon, T., Dadson, S., Buys, G. and Prudhomme, C. (2013) Bias correction of daily precipitation simulated by a regional climate model: a comparison of methods. *International Journal of Climatology*, 33(6), pp.1367-1381.
- Lempert, R.J. & Groves, D.G. (2010) Identifying and evaluating robust adaptive policy responses to climate change for water management agencies in the American west. Two Special Sections: Risk and Technology Addressing Deep Uncertainty Using Adaptive Policies, 77(6), pp.960–974.
- Liu, Y., Freer, J., Beven, K. and Matgen, P. (2009) Towards a limits of acceptability approach to the calibration of hydrological models: Extending observation error, *Journal of Hydrology*, 367(1–2), pp. 93–103.
- Liu, Z., Martina, M. L. V. & Todini, E. (2005) Flood forecasting using a fully distributed model: application of the TOPKAPI model to the Upper Xixian Catchment. *Hydrology and Earth System Sciences*, 9, 347-364.
- Lloyd-Hughes, B. & Saunders, M.A. (2002) A drought climatology for Europe. *International Journal of Climatology*, 22(13), pp.1571–1592.
- Lopez, A., Fung, F., New, M., Watts, G., Weston, A. and Wilby, R. L. (2009) From climate model ensembles to climate change impacts and adaptation: A case study of water resource management in the southwest of England, *Water Resources Research*. American Geophysical Union, 45.
- Manley, R.E. (2006) HYSIM User manual, UK: Water Resource Associates Ltd.
- Manning, L. J., Hall, J. W., Fowler, H. J., Kilsby, C. G. and Tebaldi, C. (2009) Using probabilistic climate change information from a multimodel ensemble for water resources assessment, *Water Resources Research*, 45(11), p. W11411.
- Maraun, D., Wetterhall, F., Ireson, A. M., Chandler, R. E., Kendon, E. J., Widmann, M., Brienen, S., Rust, H. W., Sauter, T., Themel, M., Venema, V. K. C., Chun, K. P., Goodess, C. M., Jones, R. G., Onof, C., Vrac, M. and Thiele-Eich, I. (2010) Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user, *Reviews of Geophysics*. AGU, 48(3), p. RG3003.
- Marsh, T, Booker, D, Fry, M. (2014) The 2004-06 Drought (2nd Edition).

- Marsh, T.J. (2002) Capitalising on river flow data to meet changing national needs -- a UK perspective. *Flow Measurement and Instrumentation*, 13(5–6).
- Marsh, T., Cole, G. & Wilby, R. (2007) Major droughts in England and Wales, 1800-2006. *Weather*, 62(4).
- Matrosov, E.S., Harou, J.J. & Loucks, D.P. (2011) A computationally efficient open-source water resource system simulator - Application to London and the Thames Basin. *Environmental Modelling and Software*, 26(12), pp.1599–1610.
- Matrosov, E.S., Woods, A.M. & Harou, J.J. (2013) Robust Decision Making and Info-Gap Decision Theory for water resource system planning. *Journal of Hydrology*, 494, pp.43–58.
- McIntyre, N., Ballard, C., Bruen, M., Bulygina, N., Buytaert, W., Cluckie, I., Dunn, S., Ehret, U., Ewen, J., Gelfan, A. and Hess, T. (2013) Modelling the hydrological impacts of rural land use change: current state of the science and future challenges. *Hydrology for a changing world*, pp.01-07.
- McKay, M. D., Conover, W. J. and Beckman, R. .J. (1979) A comparison of three methods for selection values of input variables in the analysis of output from a computer code. *Technometrics*, 2, pp.239–45.
- Mckee, T.B., Doesken, N.J. & Kleist, J. (1993) The relationship of drought frequency and duration to time scales. *AMS 8th Conference on Applied Climatology*, (January), pp.179–184.
- McMillan, H., Freer, J., Pappenberger, F., Krueger, T. and Clark, M. (2010) Impacts of uncertain river flow data on rainfall-runoff model calibration and discharge predictions. *Hydrological Processes*, 24(10), pp.1270-1284.
- McMillan, H., Krueger, T. & Freer, J. (2012) Benchmarking observational uncertainties for hydrology: rainfall, river discharge and water quality. *Hydrological Processes*, 26(26), pp.4078–4111.
- Meehl, G. A., Covey, C., Delworth, T., Latif, M., McAvaney, B., Mitchell, J., Stouffer, R. and Taylor, K. (2007) The WCRP CMIP3 multi-model dataset: A new era in climate change research, *Bulletin of the American Meteorological Society*. American Meteorological Society, 88, pp. 1383–1394.
- Met Office (2013). UKCP09: Gridded observation data sets. , 2013(September 20).

- Moore, R.J. & Clarke, R.T. (1981) A Distribution Function Approach to Rainfall Runoff Modeling. *Water Resources Research*, 17(5), pp.1367–1382.
- Moore, R.J. & Bell, V.A. (2002) Incorporation of groundwater losses and well level data in rainfall-runoff models illustrated using the PDM. *Hydrology and Earth System Sciences*, 6(1), pp.25–38.
- Moore, R. J., Bell, V. A., Cole, S. J. and Jones, D. A. (2007) Rainfall-runoff and other modelling for ungauged/low-benefit locations: operational guidelines R&D *Technical Report SC030227/SR2*. Environment Agency.
- Moore, R.J. (2007). The PDM rainfall-runoff model, *Hydrology and Earth System Sciences*, 11(1), pp.483–499.
- Moretti, G. and Montanari, A. (2007).. AFFDEF: a spatially distributed grid based rainfall–runoff model for continuous time simulations of river discharge. *Environmental Modelling & Software*, 22(6), pp.823-836.
- Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Binger, R. L., Harmel, R. D. and Veith, T. L. (2007) Model evaluation guidelines for systematic quantification of accuracy in watershed simulations, *Transactions of the ASABE*, 50(3), pp. 885–900.
- Moss, R. H., Edmonds, J. A., Hibbard, K. A., Manning, M. R., Rose, S. K., Van Vuuren, D. P., Carter, T. R., Emori, S., Kainuma, M., Kram, T. & others (2010) The next generation of scenarios for climate change research and assessment, *Nature*. Nature Publishing Group, 463(7282), pp. 747–756.
- Murphy, C., Fealy, R., Charlton, R. and Sweeney, J. (2006) The reliability of an “off-the-shelf” conceptual rainfall runoff model for use in climate impact assessment: uncertainty quantification using Latin hypercube sampling, *Area*. Irish Climate Analysis and Research Units (ICARUS), Department of Geography, NUI Maynooth, Co. Kildare, Ireland, 38(1), pp. 65–78.
- Murphy, J., Sexton, D., Jenkins, G., Boorman, P., Booth, B., Brown, K., Clark, R., Collins, M., Harris, G. and Kendon, L. (2009a) Introduction to the UKCP09 method, in *The Science of UK Climate Projection: UKCP09 and beyond*. Exeter: Royal Society.
- Murphy, J. M., Sexton, D. M. H., Jenkins, G. J., Boorman, P. M., Booth, B. B. B., Brown, C. C., Clark, R. T., Collins, M., Harris, G. R., Kendon, E. J., Betts, R. A., Brown, S. J., Howard, T., Humphrey, K. A., McCarthy, M. P., McDonald, R. E., Stephens, A., Wallace, C., Warren, R., Wilby, R. and Wood, R. A. (2009b) *UK Climate Projections Science Report: Climate change projections*. UK: Met Office.

Nakicenovic, N. & Swart, R. (2000) *IPCC Special Report on Emissions Scenarios*, UK: Cambridge Univ. Press.

Nash, J.E. & Sutcliffe, J. V. (1970) River flow forecasting through conceptual models part I: A discussion of principles. *Journal of Hydrology*, 10(3), pp.282–290. A

NERA (2012). A Non-Essential Use Drought Order for London: Economic Impact Assessment.

New, M., Lopez, A., Dessai, S. and Wilby, R. (2007a) Challenges in using probabilistic climate change information for impact assessments: an example from the water sector, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365(1857), pp. 2117–2131.

New, M., Cuellar, M. & Lopez, A. (2007b) Probabilistic regional and local climate projections: false dawn for impacts assessment and adaptation?

Newton, G., Dadson, S., Lafon, T. and Prudhomme, C. (2012) Scoping Study for Precipitation Downscaling and Bias-Correction. *Science Report/Project Note - SC090016/PN3*. DEFRA.

Norton, T., Byrne, T. (2004) Modelling the Hydrological Impacts of Climate Change on Southern Region River Flows. Lead Area Hydrology, Southern Region: UK Environment Agency.

Ntegeka, V., Baguis, P., Roulin, E. and Willems, P. (2014) Developing tailored climate change scenarios for hydrological impact assessments, *Journal of Hydrology*. Elsevier, 508, pp. 307–321.

Office for National Statistics (2014). 2012-based National Population Projections, 1951-2087.

Office for National Statistics, ONS, (2013a). 2012-based National Population Projections Low Fertility variant projection, England, 2012-2112.

Office for National Statistics, ONS, (2013b). Office for National 2012-based National Population Projections High Fertility variant projection, England, 2012-2112.

Oni, S., Futter, M., Ledesma, J., Teutschbein, C., Buttle, J. and Laudon, H. (2016) Using dry and wet year hydroclimatic extremes to guide future hydrologic projections. *Hydrology and Earth System Sciences*, 20(7), pp.2811-2825.

Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andrassian, V., Anctil, F. and Loumagne, C. (2005) Which potential evapotranspiration input for a lumped rainfall-runoff model?: Part 2: Towards a

simple and efficient potential evapotranspiration model for rainfall-runoff modelling, *Journal of Hydrology*, 303(1–4), pp. 290–306.

Page, T., Beven, K. J., Freer, J. and Jenkins, A. (2003) Investigating the uncertainty in predicting responses to atmospheric deposition using the model of acidification of groundwater in catchments (MAGIC) within a generalised likelihood uncertainty estimation (GLUE) framework, *Water, air, and soil pollution*. Springer, 142(1–4), pp. 71–94.

Parker, J.M. (2014) Assessing the sensitivity of historic micro-component household water-use to climatic drivers. A Doctoral Thesis. Submitted in partial fulfilment of the requirements for the award of Doctor of Philosophy of Loughborough University, Loughborough: Loughborough. <https://dspace.lboro.ac.uk/2134/14939>.

Parry, S., Wilby, R.L., Prudhomme, C. and Wood, P.J. (2016a) A systematic assessment of drought termination in the United Kingdom. *Hydrology and Earth System Sciences*, 20(10), p.4265.

Parry, S., Prudhomme, C., Wilby, R.L. and Wood, P.J. (2016b) Drought termination: Concept and characterisation. *Progress in Physical Geography*, 40(6), pp.743-767.

Penman, H.L. (1948) Natural evaporation from open water, bare soil and grass. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*. pp. 120–145.

Peters, G. P., Andrew, R. M., Boden, T., Canadell, J. G., Ciais, P., Le Quéré, C., Marland, G., Raupach, M. R. and Wilson, C. (2013) The challenge to keep global warming below 2 C, *Nature Climate Change*. Nature Publishing Group, 3(1), pp. 4–6.

Piani, C. & Haerter, J.O. (2012) Two dimensional bias correction of temperature and precipitation copulas in climate models. *Geophysical Research Letters*, 39(20).

Pianosi, F. & Wagener, T. (2016) Understanding the time-varying importance of different uncertainty sources in hydrological modelling using global sensitivity analysis. *Hydrological Processes*.

Prudhomme, C., Piper, B., Osborn, T. and Davies, H. (2005) *Climate Change Uncertainty in Water Resource Planning*. UK: UKWIR.

Prudhomme, C., Haxton, T., Crooks, S., Jackson, C., Barkwith, A., Williamson, J., Kelvin, J., Mackay, J., Wang, L., Young, A. and Watts, G. (2012) *Future Flows Hydrology: an ensemble of daily river*

flow and monthly groundwater levels for use for climate change impact assessment across Great Britain, *Earth System Science Data Discussions*, 5(2), pp. 1159–1178.

Prudhomme, C., Wilby, R. L., Crooks, S., Kay, A. L. and Reynard, N. S. (2010) Scenario-neutral approach to climate change impact studies: Application to flood risk, *Journal of Hydrology*, 390(3–4), pp. 198–209.

Prudhomme, C. & Davies, H. (2009). Assessing uncertainties in climate change impact analyses on the river flow regimes in the UK Part 2: future climate. *Climatic Change*, 93(1).

Prudhomme, C. & Williamson, J. (2013) Derivation of RCM-driven potential evapotranspiration for hydrological climate change impact analysis in Great Britain: a comparison of methods and associated uncertainty in future projections. *Hydrology and Earth System Sciences*, 17(4), pp.1365–1377.

Pushpalatha, R., Perrin, C., Moine, N. Le and Andrassian, V. (2012) A review of efficiency criteria suitable for evaluating low-flow simulations, *Journal of Hydrology*, 420–421(0), pp. 171–182.

Rance, J., Wade, S.D., Hurford, A.P., Bottius, E. and Reynard, N.S. (2012) CCRA Risk Assessment for the Water Sector. *UK 2012 Climate Change Risk Assessment*.

Refsgaard, J.C., Stomr, B. & Abbott, M.B. (1996) Comment on: A discussion of distributed hydrological modeling by K Beven. *Distributed hydrological modeling*, pp.279–287.

Reynard, N. S., Crooks, S. M., Kay, A. L. and Prudhomme, C. (2010) *Regionalised impacts of climate change on flood flows*. UK: DEFRA.

Robinson, E.L.; Blyth, E.; Clark, D.B.; Comyn-Platt, E.; Finch, J.; Rudd, A.C. (2016) Climate hydrology and ecology research support system potential evapotranspiration dataset for Great Britain (1961-2015) [CHESS-PE] . NERC Environmental Information Data Centre.

Rocheta, E., Sugiyanto, M., Johnson, F., Evans, J. and Sharma, A. (2014) How well do general circulation models represent low-frequency rainfall variability?, *Water Resources Research*, 50(3), pp. 2108–2123.

Rushton, K.R., Eilers, V.H.M. and Carter, R.C. (2006) Improved soil moisture balance methodology for recharge estimation. *Journal of Hydrology*, 318(1-4), pp.379-399.

Sahu, R.K., Mishra, S.K. & Eldho, T.I. (2010) An improved AMC-coupled runoff curve number model. *Hydrological Processes*, 24(20), pp.2834–2839.

- Scaife, A. A., Woollings, T., Knight, J., Martin, G. and Hinton, T. (2010) Atmospheric blocking and mean biases in climate models, *Journal of Climate*, 23(23), pp. 6143–6152.
- Scaife, A.A., Copsey, D., Gordon, C., Harris, C., Hinton, T., Keeley, S., O'Neill, A., Roberts, M. and Williams, K. (2011) Improved Atlantic winter blocking in a climate model. *Geophysical Research Letters*, 38(23).
- Schewe, J., Heinke, J., Gerten, D., Haddeland, I., Arnell, N.W., Clark, D.B., Dankers, R., Eisner, S., Fekete, B.M., Colón-González, F.J. and Gosling, S.N. (2014) Multimodel assessment of water scarcity under climate change. *Proceedings of the National Academy of Sciences*, 111(9), pp.3245-3250.
- Schulla, J. and Jasper, K. (2007) Model description Wasim-eth. Institute for Atmospheric and Climate Science, Swiss Federal Institute of Technology, Zürich.
- Seiller, G. & Anctil, F. (2016) How do potential evapotranspiration formulas influence hydrological projections? *Hydrological Sciences Journal*, 61(12), pp.2249–2266.
- Sexton, D. H., Murphy, J., Collins, M. and Webb, M. (2012) Multivariate probabilistic projections using imperfect climate models part I: outline of methodology, *Climate Dynamics*. Springer-Verlag, 38(11–12), pp. 2513–2542. doi: 10.1007/s00382-011-1208-9.
- Sexton, D.M. and Harris, G.R. (2015) The importance of including variability in climate change projections used for adaptation. *Nature Climate Change*, 5(10), pp.931-936.
- Shiklomanov, I., (1993) World fresh water resources in Gleick, P.H.: *Water in crisis: a guide to the worlds fresh water resources*.
- Silburn, D.M. & Connolly, R.D. (1995) Distributed parameter hydrology model (ANSWERS) applied to a range of catchment scales using rainfall simulator data I: Infiltration modelling and parameter measurement. *Journal of Hydrology*, 172(1–4).
- Sillmann, J., Thorarinsdottir, T., Keenlyside, N., Schaller, N., Alexander, L.V., Hegerl, G., Seneviratne, S.I., Vautard, R., Zhang, X. and Zwiers, F.W. (2017) Understanding, modeling and predicting weather and climate extremes: Challenges and opportunities. *Weather and climate extremes*, 18, pp.65-74.
- Sivapalan, M. (2005) Pattern, process and function: elements of a unified theory of hydrology at the catchment scale. *Encyclopedia of hydrological sciences*.



- Smith, M. B., Seo, D.-J., Koren, V. I., Reed, S. M., Zhang, Z., Duan, Q., Moreda, F. and Cong, S. (2004) The distributed model intercomparison project (DMIP): motivation and experiment design, *Journal of Hydrology*, 298(1–4).
- Smith, A., Freer, J., Bates, P. and Sampson, C. (2014) Comparing ensemble projections of flooding against flood estimation by continuous simulation. *Journal of Hydrology*, 511, pp.205–219.
- Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Avery, K. B., Tignor, M. and Miller, H. L. (2007) *Climate Change 2007 - The Physical Science Basis: Working Group I Contribution to the Fourth Assessment Report of the IPCC*. Edited by H. L. Miller. Cambridge, UK and New York, NY, USA: Cambridge University Press.
- Southern Water (2014). Water resources management plan 2015-2040.
- Spear, R.C. & Hornberger, G.M. (1980) Eutrophication in Peel Inlet:II. Identification of critical uncertainties via generalized sensitivity analysis. *Water Research*, 14(1), pp.43–49.
- Stagge, J. & Tallaksen, L. (2014) Standardized precipitation-evapotranspiration index (SPEI): Sensitivity to potential evapotranspiration model and parameters. *International Association of Hydrological Sciences (IAHS)*, 10(October), pp.367–373.
- Stagge, J. H., Tallaksen, L. M., Gudmundsson, L., Van Loon, A. F. and Stahl, K. (2015) Candidate Distributions for Climatological Drought Indices (SPI and SPEI), *International Journal of Climatology*, 35(13), pp. 4027–4040. doi: 10.1002/joc.4267.
- Stahl, K., Hisdal, H., Hannaford, J., Tallaksen, L., Van Lanen, H., Sauquet, E., Demuth, S., Fendekova, M. and Jordar, J. (2010) Streamflow trends in Europe: evidence from a dataset of near-natural catchments, *Hydrology and Earth System Sciences*, 14, p. p--2367.
- Stainforth, D. A., Downing, T. E., Washington, R., Lopez, A. and New, M. (2007) Issues in the interpretation of climate model ensembles to inform decisions, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365(1857).
- Stainforth, D. A., Aina, T., Christensen, C., Collins, M., Faull, N., Frame, D. J., Kettleborough, J. A., Knight, S., Martin, A., Murphy, J. M., Piani, C., Sexton, D., Smith, L. A., Spicer, R. A., Thorpe, A. J. and Allen, M. R. (2005) Uncertainty in predictions of the climate response to rising levels of greenhouse gases, *Nature*, 433(7024).

- Stedinger, J. R., Vogel, R. M., Lee, S. U. and Batchelder, R. (2008) Appraisal of the generalized likelihood uncertainty estimation (GLUE) method, *Water Resources Research*. AGU, 44, p. W00B06.
- Steele-Dunne, S., Lynch, P., McGrath, R., Semmler, T., Wang, S., Hanafin, J. and Nolan, P. (2008) The impacts of climate change on hydrology in Ireland, *Journal of Hydrology*, 356(1–2), pp. 28–45.
- Tanguy, M., Hannaford, J., Barker, L., Svensson, C., Kral, F., Fry, M. (2015) Gridded Standardised Precipitation-Evapotranspiration Index (SPEI) using generalised logistic distribution with standard period 1961-2010 for Great Britain.
- Tanguy, M.; Dixon, H.; Prosdocimi, I.; Morris, D. G.; Keller, V. D. J. (2016) Gridded estimates of daily and monthly areal rainfall for the United Kingdom (1890-2015) [CEH-GEAR]. NERC Environmental Information Data Centre.
- Teutschbein, C., Seibert, J. (2013) "Is bias correction of regional climate model (RCM) simulations possible for non-stationary conditions?." *Hydrology and Earth System Sciences* 17.12: 5061.
- Thames Water (2010). Thames Water Drought Plan.
- Thames Water (2012). Revised Estimated Impacts of Demand-side drought management activities.
- Thames Water (2013). Final Drought Plan.
- Thames Water (2014). Thames Water Final Water Resources Management Plan 2014.
- Thames Water (2018). Thames Water draft Water Resources Management Plan 2019.
- Tian, Y., Booij, M.J. and Xu, Y.P. (2014) Uncertainty in high and low flows due to model structure and parameter errors. *Stochastic environmental research and risk assessment*, 28(2), pp.319-332.
- Todini, E. (1996) The ARNO rainfall-runoff model. *Journal of Hydrology*, 175(1–4), pp.339–382.
- Todini, E. and Ciarapica, L. (2002) The TOPKAPI model. *Mathematical models of large watershed hydrology*, pp.471-506.
- Turner, S. W. D., Marlow, D., Ekstr, M. and Jeffrey, P. J. (2014) Linking climate projections to performance: A yield-based decision scaling assessment of a large urban water resources system, *Water Resources Research*, 50(4), pp. 3553–3567.

- Uhlenbrook, S., Seibert, J. A. N., Leibundgut, C. and Rodhe, A. (1999) Prediction uncertainty of conceptual rainfall-runoff models caused by problems in identifying model parameters and structure, *Hydrological Sciences Journal*. Taylor & Francis, 44(5), pp. 779–797.
- Uhlenbrook, S., Roser, S. & Tilch, N. (2004) Hydrological process representation at the meso-scale: the potential of a distributed, conceptual catchment model. *Journal of Hydrology*, 291(3–4), pp.278–296.
- UKCP09 (2017). UK Climate Projections. Wind or Wind speed.
- UKCIP (2010). UK Climate Impacts Programme.
- UKWIR (2007). Effects of Climate Change on River Flows and Groundwater Recharge, a Practical Methodology: Synthesis Report, *UKWIR Report 07/CL/04/10*, London.
- UKWIR (2012). Impact of climate change on water demand. *UKWIR Report CL04B*, London.
- UKWIR (2016a). Climate Change Modelling & the WRMP: Project Resume. *UKWIR Project CL04*, London.
- UKWIR (2016b). Water Resources Planning Methods 2019 –Decision making process : guidance.
- UKWIR (2016c). Water Resources Planning Methods 2019 Methods – Risk based planning.
- Vaze, J., Post, D.A., Chiew, F.H.S., Perraud, J.M., Viney, N.R. and Teng, J. (2010) Climate non-stationarity–validity of calibrated rainfall–runoff models for use in climate change studies. *Journal of Hydrology*, 394(3-4), pp.447-457.
- Velázquez, J.A., Schmid, J., Ricard, S., Muerth, M.J., St-Denis, B.G., Minville, M., Chaumont, D., Caya, D., Ludwig, R. and Turcotte, R. (2013) An ensemble approach to assess hydrological models' contribution to uncertainties in the analysis of climate change impact on water resources. *Hydrology and Earth System Sciences*, 17, pp.565-578.
- Vicente-Serrano, S.M., Beguera, S. & Lopez-Moreno, J.I. (2010) A multi-scalar drought index sensitive to global warming: The standardized precipitation evapotranspiration index. *Journal of Climate*, 23(7), pp.1696–1718.
- Vicente-Serrano, S.M. & Beguería, S. (2016) Comment on ‘Candidate distributions for climatological drought indices (SPI and SPEI)’by James H. Stagge et al. *International Journal of Climatology*, 36(4), pp.2120-2131.

- Wagener, T., Boyle, D.P., Lees, M.J., Wheater, H.S., Gupta, H.V. and Sorooshian, S. (2001) A framework for development and application of hydrological models. *Hydrology and Earth System Sciences Discussions*, 5(1), pp.13-26.
- Wagener, T., Wheater, H. & Gupta, H.V. (2004) Rainfall-runoff modelling in gauged and ungauged catchments, Imperial College Press.
- Watts, G., Battarbee, R.W., Bloomfield, J.P., Crossman, J., Daccache, A., Durance, I., Elliott, J.A., Garner, G., Hannaford, J., Hannah, D.M. and Hess, T. (2015) Climate change and water in the UK—past changes and future prospects. *Progress in Physical Geography*, 39(1), pp.6-28.
- Walsh, C.L., Blenkinsop, S., Fowler, H.J., Burton, A., Dawson, R.J., Glenis, V., Manning, L.J. and Kilsby, C.G. (2015) Adaptation of water resource systems to an uncertain future. *Hydrology and Earth System Sciences Discussions*, 12(9), pp.8853-8889.
- Water Act (2010). Water Act 2003. Statute.
- Water Resources Act (1991). Water Resources Act 1991. Statute.
- Westerberg, I.K. & McMillan, H.K. (2015) Uncertainty in hydrological signatures. *Hydrology and Earth System Sciences*, 19(9), pp.3951-3968.
- Westerberg, I.K., Guerrero, J.L., Younger, P.M., Beven, K.J., Seibert, J., Halldin, S., Freer, J.E. and Xu, C.Y. (2011a) Calibration of hydrological models using flow-duration curves. *Hydrology and Earth System Sciences*, 15(7), p.2205.
- Westerberg, I., Guerrero, J.L., Seibert, J., Beven, K.J. and Halldin, S. (2011b) Stage-discharge uncertainty derived with a non-stationary rating curve in the Choluteca River, Honduras. *Hydrological Processes*, 25(4), pp.603-613.
- Westerberg, I. K., Wagener, T., Coxon, G., McMillan, H. K., Castellarin, A., Montanari, A. and Freer, J. (2016) Uncertainty in hydrological signatures for gauged and ungauged catchments, *Water Resources Research*, 52(3), pp. 1847–1865.
- Whetton, P., Hennessy, K., Clarke, J., McInnes, K. and Kent, D. (2012) Use of Representative Climate Futures in impact and adaptation assessment. *Climatic Change*, 115(3-4), pp.433-442.
- White, E.R., Easton, Z.M., Fuka, D.R. and Steenhuis, T.S. (2009) SWAT-WB theoretical documentation. Soil and Water Lab, Department of biological and Environmental engineering, Cornell University, Ithaca, NY.

- Whitehead, P.G., Wilby, R.L., Battarbee, R.W., Kernan, M. and Wade, A.J. (2009) A review of the potential impacts of climate change on surface water quality. *Hydrological Sciences Journal*, 54(1), pp.101-123.
- Wilby, R., Greenfield, B. & Glenny, C. (1994) A coupled synoptic-hydrological model for climate change impact assessment. *Journal of Hydrology*, 153(1–4), pp.265–290.
- Wilby, R.L. (2005). Uncertainty in water resource model parameters used for climate change impact assessment. *Hydrological Processes*, 19(16), pp.3201–3219.
- Wilby, R.L. & Harris, I., (2006) A framework for assessing uncertainties in climate change impacts: Low-flow scenarios for the River Thames, UK. *Water Resources Research.*, 42.
- Wilby, R. L., Troni, J., Biot, Y., Tedd, L., Hewitson, B. C., Smith, D. M. and Sutton, R. T. (2009) A review of climate risk information for adaptation and development planning, *International Journal of Climatology*, 29(9), pp. 1193–1215.
- Wilby, R.L., Fenn, C.R., Wood, P.J., Timlett, R. and LeQuesne, T. (2011) Smart licensing and environmental flows: Modeling framework and sensitivity testing. *Water Resources Research*, 47(12).
- Willems, P., Arnbjerg-Nielsen, K., Olsson, J. and Nguyen, V. T. V. (2012) Climate change impact assessment on urban rainfall extremes and urban drainage: Methods and shortcomings, *Atmospheric Research*, 103, pp. 106–118.
- Wood, A. W., Leung, L. R., Sridhar, V. and Lettenmaier, D. P. (2004) Hydrologic Implications of Dynamical and Statistical Approaches to Downscaling Climate Model Outputs, *Climatic Change*, 62(1), pp. 189–216.
- World Meteorological Organization (2012). Standardized Precipitation Index User Guide (WMO-No. 1090), p.24.
- Yates, D.N., Miller, K.A., Wilby, R.L. and Kaatz, L. (2015) Decision-centric adaptation appraisal for water management across Colorado’s continental divide. *Climate Risk Management*, 10, pp.35-50.
- Young, P.C. & Beven, K.J. (1994) Data-based mechanistic modelling and the rainfall-flow non-linearity. *Environmetrics*, 5(3), pp.335-363.

Young, P.C. (2001) Data-based mechanistic modelling and validation of rainfall-flow processes. In M. G. Anderson & P. D. Bates, eds. *Model Validation: Perspectives in Hydrological Sciences*. Chichester: Wiley, pp. 117–161.

Young, P.C. & Ratto, M. (2009) A unified approach to environmental systems modeling. *Stochastic Environmental Research and Risk Assessment*, 23(7), pp.1037-1057.

Young, P.C. (2013) Hypothetico-inductive data-based mechanistic modeling of hydrological systems. *Water Resources Research*, 49(2), pp.915-935.

Yu, P.S., Yang, T.C. & Chen, S.J. (2001) Comparison of uncertainty analysis methods for a distributed rainfall-runoff model. *Journal of Hydrology*, 244(1–2), pp.43–59.

Zhao, L., Xia, J., Xu, C.Y., Wang, Z., Sobkowiak, L. and Long, C. (2013). Evapotranspiration estimation methods in hydrological models. *Journal of Geographical Sciences*, 23(2), pp.359-369.



## **Appendix A – Inventory of selected conceptual hydrological models**





Name	Model variants	Model type	Run-off generation	Run-off routing
AFFDEF (Moretti & Montanari, 2007)	None	Semi-distributed (grid) conceptual model	Semi-empirical (based on modified Curve Number with linear stores/reservoirs)	Surface (hillslope and river) and sub-surface flow via Muskingum-Cunge model (see Chow et al., 1988)
ARNO (Todini, 1996)	None	Semi-distributed (distribution function) conceptual model	Variable area saturation excess with a soil moisture balance based on a probability distribution function	Hillslope and channel routing using distributed inflow linear parabolic approaches
CatchMod (Wilby et al., 1994)	None	Lumped or semi-distributed (HRUs) conceptual model	Fixed area, saturation excess (with uniform two layer soil moisture store) with optional fixed percentage direct bypass	Two sequential stores (linear to non-linear) and optional open-channel routing
CLASSIC (Crooks & Naden, 2007)	None	Semi-distributed (grid), conceptual model	Saturation excess with two parameter soil moisture balance model. Urban areas with separate soil moisture balance	Drainage model with one or two (parallel) linear stores plus urban areas. Basin wide channel routing via two parameter model (wave velocity and coefficient of diffusion).
Grid-to-Grid	Slope-G2G (Bell et al., 2007a)	Semi-distributed (grid) conceptual model	Saturation excess based on probability distributed soil moisture store. Non-linear (typically cubic) storage function relates groundwater flow to volume	Parallel coupled equations based on 1-dimensional kinematic wave for surface and sub-surface flow. Channel flow also based on channel width and slope
	Soil-G2G (Bell et al., 2009)		As for Slope-G2G but storage based on soil data. Non-linear (typically cubic) storage function relates groundwater flow to volume	
GRID Model (Bell & Moore, 1998)	None	Semi-distributed (grid) conceptual model	Saturation excess based on probability distributed soil moisture store or topographic index	Routed from cell to outlet (not square to square) based on DTM-derived isochrone pathways
HBV (Bergström, 1995)	HBV-96	Lumped or semi-distributed (HRUs) conceptual model	Saturation excess using soil moisture balance with statistical distribution of storage capacities	Linear (baseflow) and non-linear (fast) reservoirs with sub-basins linked via simplified Muskingum approach or time lag
	HBV-Light	Lumped or semi-distributed (HRUs) conceptual model		3 linear reservoir equations with time lag

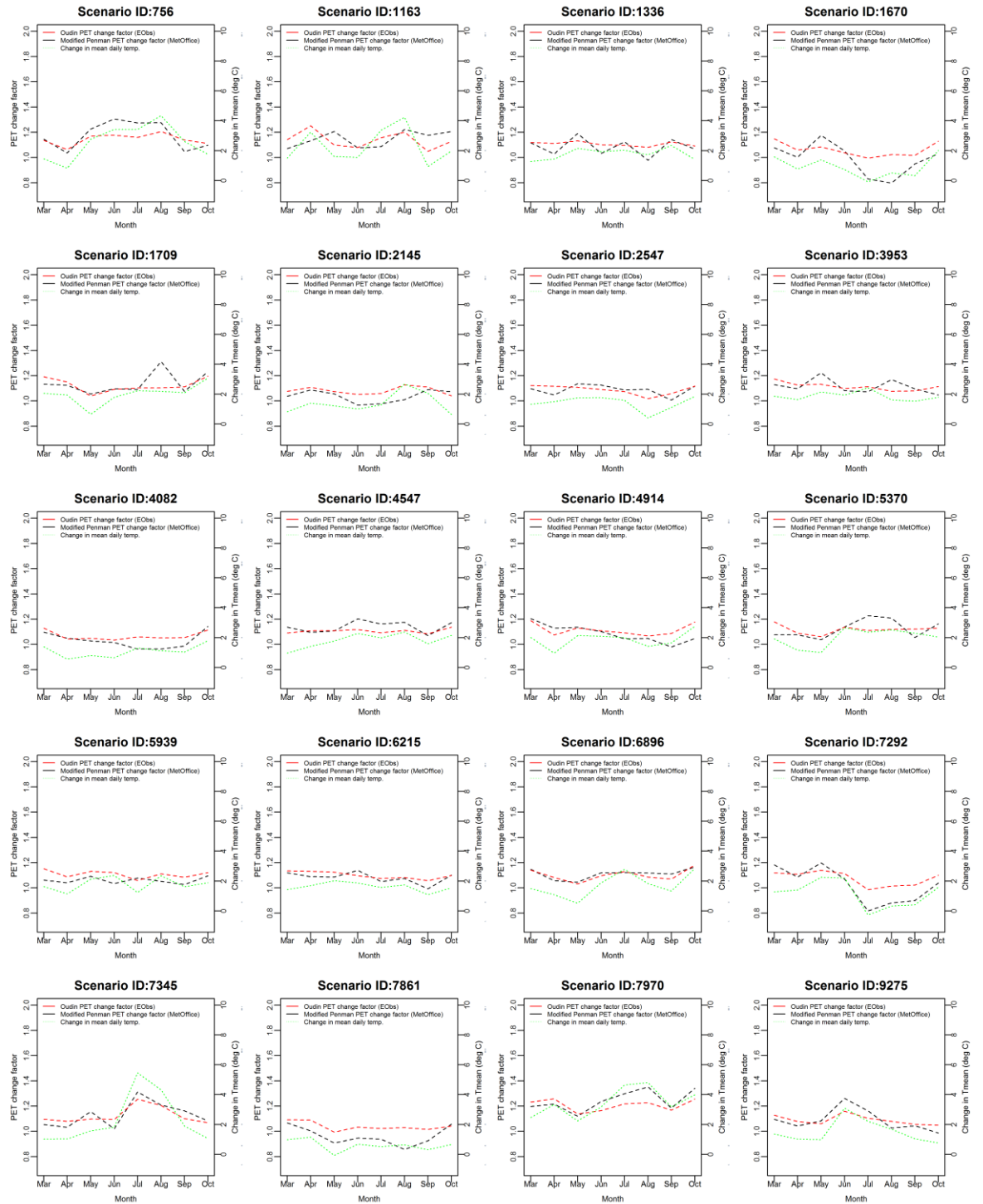
Name	Model variants	Model type	Run-off generation	Run-off routing
HEC-HMS (Chu & Steinman, 2009)	None	Lumped or semi-distributed (HRU or gridded) conceptual model	Infiltration losses calculated using several methods including a single layer soil moisture store or a five layer soil moisture store. Linear reservoirs can be used to link infiltration losses to baseflow	Unit hydrograph approaches or kinematic wave. Hydrological (reach) routing methods include Muskingum, Muskingum-Cunge, time lag, Kinematic wave and modified Puls methods (see Chow et al., 1988)
HSPF (Johnson et al., 2003)	None	Semi-distributed (HRUs) physically-based model	Infiltration-excess model that separates moisture inputs into infiltrating and non-infiltrating fractions.. A linear probability distribution is used to reflect areal variation in infiltration capacity	Overland flow routed using Chezy–Manning equation (see Chow et al., 1988). Channel routing using kinematic wave method
HYSIM (Manley, 2006)	None	Semi-distributed (HRUs) conceptual model	Infiltration excess (Philips equation) with two layer soil moisture store. Two linear reservoirs used for groundwater	Kinematic flow routing (see Chow et al., 1988)
PDM (Moore, 2007)	Multiple variants – see main text	Lumped or semi-distributed (distribution function) conceptual model	Saturation excess with a single store soil moisture balance with a probability distribution of soil moisture storage	Two parallel (fast and slow) routing processes based on a variety of non-linear storage reservoirs or by a cascade of two linear reservoirs
SMDR (Johnson et al., 2003)	None	Semi-distributed (grid) conceptual model	Saturation excess with a soil moisture balance with up to four stores. Conceptual lumped aquifer store with outflow based on TOPMODEL approach	Subsurface lateral flow routed using Darcy's equation assuming hydraulic gradient parallel to slope
SWAT based models (SWAT, SWAT-VSA and SWAT-WB)	SWAT / SWAT-CN (Arnold et al., 1998)	Semi-distributed (HRUs) empirical/conceptual model	Empirical (Curve Number) or Infiltration Excess (Green & Ampt). Bypass flows can also be specified. Kinematic storage model for subsurface flow	Time lag for run-off
	SWAT-VSA		Saturation excess - adjustment of Curve Number based on topographic index	Channel routing using Kinematic wave (Muskingum or variable storage variants)
	SWAT-WB (White, 2009)		Saturation excess using a soil moisture balance	

Name	Model variants	Model type	Run-off generation	Run-off routing
TOPKAPI (Ciarapica & Todini, 2002)	None	Semi-distributed (grid) conceptual model	Saturation excess soil moisture balance	Overland, sub-surface and channel flow routed by kinematic wave approximation (Muskingum-Cunge, see Chow et al., 1988)
TOPMODEL	Original version (Beven, 2001)	Semi-distributed (topographic index) conceptual model	Saturation excess with storage based on topographic index. Two stores to model interception/rootzone and drainage/recharge.	Overland flow delay timestep for larger catchments. Channel routing typically based on simple constant speed routing
	Dynamic TOPMO DEL (Beven & Freer, 2001a)	Semi-distributed (topographic index) conceptual model	As TOPMODEL but allowing for dynamically varying upslope contributing areas in the calculation of the topographic index	
WASIM-ETH (Schulla & Jasper, 2007)	Version 1 (TOPMO DEL)	Semi-distributed (grid) conceptual model	Grid cell based modified variable saturated area approach (similar to TOPMODEL topographic index approach)	Kinematic wave approach (translation) with linear reservoir (attenuation and diffusion)
	Version 2 (Richards equation)		Infiltration excess based on Green and Ampt equation. Vertical flow in unsaturated soil based on Richard's equation. Integrated 2D-groundwater model available	
WetSpa (Bahremand et al., 2007)	WetSpa-ss (steady state) also available	Semi-distributed (gridded or HRUs) conceptual /empirical model	Rainfall excess based on a moisture-related modified rational method with a potential runoff coefficient depending on the land cover, soil type, slope, magnitude of rainfall, and the antecedent soil moisture. Interflow computed using Darcy's Law. Linear or non-linear reservoirs for groundwater flow with lumped store	Linear diffusive wave approximation for overland and channel flow



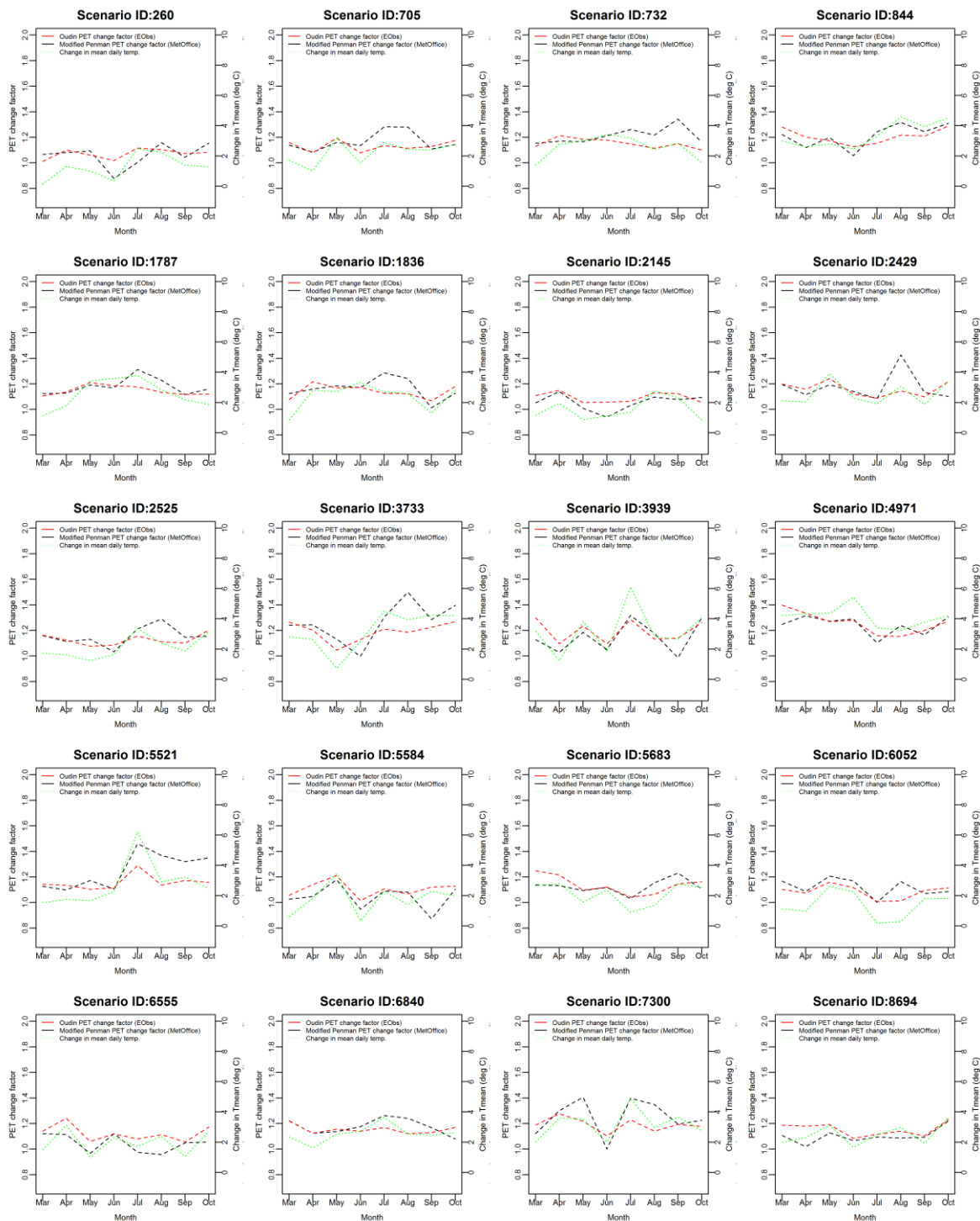
## **Appendix B – Mean monthly changes in PET for the 2030s and 2050s**





**Figure B. 1: Comparison of mean monthly changes in PET under the Oudin (red dashed lines) and modified Penman-Monteith (black dashed lines) methods and the mean monthly change in mean daily temperature (green dashed lines) for 20 UKCP09 sampled projections for the 2030s under a Medium Emission scenario. Changes are relative to a 1961 to 1990 baseline.**





**Figure B. 2: Comparison of mean monthly changes in PET under the Oudin (red dashed lines) and modified Penman-Monteith (black dashed lines) methods and the mean monthly change in mean daily temperature (green dashed lines) for 20 UKCP09 sampled projections for the 2050s under a Medium Emission scenario. Changes are relative to a 1961 to 1990 baseline.**